

# nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE



## MAKING CONNECTIONS

Lost neurons replaced by  
embryonic cells in mouse

neocortex **PAGE 248**

### NEURODEGENERATION

#### THE DEMENTIA TIME BOMB

We need more money, better  
therapeutics — and luck

**PAGES 156 & 166**

### BRAIN RESEARCH

#### GRASSROOTS CAMPAIGN

Bottom-up collaborations  
share goals, data, methods

**PAGE 159**

### NANOTECHNOLOGY

#### MOLECULAR ACTION REPLAY

Frame-by-frame imaging  
of single molecules

**PAGES 170 & 263**

**NATURE.COM/NATURE**

10 November 2016 £10

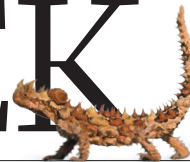
Vol. 539, No. 7628



9 770028 083095



# THIS WEEK



## EDITORIALS

**WRITING** Tracking the shifting norms of academic language **p.140**

**WORLD VIEW** The insect extinction crisis we can't ignore **p.141**

**DRINK UP** Thorny devil could drain water from wet sand **p.142**

## Environmental rights

*Brazil is suffering from both an economic and a political crisis, but eliminating basic environmental protections is no solution.*

Brazil has had its ups and downs when it comes to protecting the environment, but on paper, at least, many of the country's policies are admirably green. The right to an "ecologically balanced environment" is even enshrined in the Brazilian constitution. Now, however, a loose-knit coalition of agricultural and industrial interests is working to undermine the government's authority — and constitutional obligation — to protect the environment.

At issue is Brazil's environmental-licensing system, which governs infrastructure projects ranging from petrol stations to ports, dams and mines. Following international norms, the Brazilian environment ministry reviews and assesses such projects to ensure that they follow the law and protect the environment without infringing on the rights of local communities. This is a cornerstone of modern environmental regulation, but pro-business lawmakers are concerned that it is getting in the way of progress.

As discussed in a News story on page 147, a variety of proposals to streamline the process have been floated in the Brazilian Congress. All are headed in the wrong direction.

The debate has simmered for years, but conservative lawmakers are now capitalizing on the economic recession, corruption scandals and political turmoil that have rocked the country in recent years. Many of the same business interests were behind the 2012 law that weakened protections under Brazil's 1965 Forest Code, a landmark piece of environmental legislation that governs forested lands across the country.

Together, these efforts mark a backlash against the regulatory efforts that helped Brazil to slash the rate of deforestation to a historic low of 4,571 square kilometres in 2012. Since then, however, it has gone up by more than one-third, and could go higher when Brazil releases the numbers for 2016 in the coming days.

As it stands, the Brazilian government has a national plan for energy infrastructure that extends to 2030 and is heavily weighted towards hydroelectricity. The problem is that the plan was apparently produced with little public input, and contains only a simplistic assessment of the environmental and social impacts of installing dams in the Amazon. Once dams are formally proposed, they hit a wall of public opposition.

One solution is to bolster public participation and environmental review during such strategic planning processes. This would enable a broader dialogue among communities, indigenous groups, companies and government officials about where such projects can be placed with the least environmental and social impacts. The process could also focus on the cumulative impacts across the Amazon biome, rather than just the local effects of a particular project. This would take time and resources, but it might head off some of the protests and legal challenges that afflict so many projects today.

The idea of instituting strategic environmental assessments was included in early drafts of environmental-licensing legislation being developed by the environment ministry. This would be a step in the right direction. And if Brazil looked broadly at energy options, it might

also discover that wind and solar electricity offer better opportunities in many places around the country, with fewer risks and headaches. Hydropower has helped Brazil to maintain a low-carbon footprint up until now, but the country should be wary of betting its future on

rainfall that may shift to another region as the planet warms in the coming decades.

**"Partnerships helped to drive down deforestation, and everybody benefited."**

Over the past decade, major beef and soybean exporters have made commitments to ensure that they are not contributing to deforestation in Brazil, including by signing moratoria on the purchase of products that come from recently cleared land. Those

partnerships among businesses, environmentalists and governments helped to drive down deforestation, and everybody benefited. Today, the powerful agricultural sector continues to grow, despite the recession. But Brazil's reputation, as well as that of its most powerful industry, is back in the balance. ■

## Set prejudice aside

*Fair evaluation of science requires that the work is judged on its merits, not on assumptions.*

In April, for the first time, a couple gave birth to a healthy baby conceived using a new technique that mixes the DNA from three people. This mitochondrial replacement technology seems to have eliminated the disease, the group that performed the procedure announced in September. The scientists and clinicians at the New Hope Fertility Center in Mexico were proud that it happened in their country.

That pride soon turned sour. Scientists and ethics societies elsewhere — who have spent years drawing up guidelines for such a procedure — responded to the surprise announcement with criticism. In interviews and at meetings, researchers and experts raised vague doubts about whether the New Hope team had properly informed their patients, or whether it had broken laws. They pointed out the number of back-alley, questionable stem-cell treatment clinics that exist in Mexico. The implication was clear: the group that did the research had not played by the proper rules, by which they meant the rules the experts themselves had agreed.

New Hope's clinic in Guadalajara is inspected for quality by federal regulators, and the researchers say that an institutional review board (IRB) had approved the mitochondrial replacement project in accordance with federal law. "Why is an IRB in the UK better than ours in Mexico?" medical director Alejandro Chávez-Badiola asked



*Nature* in an interview. “This is scientific chauvinism.”

In fairness, the bad impression was exacerbated by the New Hope clinic’s decision to announce its news through the popular media — complete with a rash comment about the lack of research scrutiny — and at a scientific meeting, rather than publishing in a journal. It is true, for example, that the procedure is impossible under current US stipulations. But insinuations that research is problematic purely because of where it is performed are outdated, damaging and elitist.

Stereotypes are not uncommon when researchers talk about those in a different country. They came into focus last year, for instance, when a group of Chinese scientists edited the DNA of a human embryo for the first time. Most reporters who asked around will have heard the same condescending statements about the supposed lower ethical standards in China. It is hard to believe that the same scientists do not make the same comments to colleagues and others.

History can introduce a similar bias. Countries that were denied access to antibiotics for decades during the cold war, such as Georgia, have years of data on the efficacy of phage therapy — killing bacteria with viruses found in the environment. The idea should work, but some infectious-disease specialists in countries such as the United States and the United Kingdom hesitate to consider the potential of phage therapy, largely because it is something that is done in former Soviet states and not in ‘modern’ medicine. So the cycle becomes self-sustaining, even in the face of rampant antibiotic resistance.

This is not the sort of criticism that shows up in opinion articles or in letters to journals. It is insidious and therefore hard to confront directly. But neither is it necessarily malicious or even intentional. Like any microaggression, it’s more likely to be implied, in meaningful statements along the lines of, “Well, it was in Russia, after all...”. It is dropped into casual conversation with colleagues or mentioned to reporters by way of explanation. It is impossible to challenge precisely because it is vague and implies ‘they don’t do things our way’.

There is, of course, a level of realism that should apply to the capabilities of any single group. Discussions need to be had about whether international accord should be reached in certain areas, such as the inviolable Declaration of Helsinki rules on human experimentation. And researchers should not shy away from demanding that researchers from any country be open about the ethical and scientific underpinnings of their work.

But assumptions writ large about science in a particular country harm relationships and risk creating a backlash. They are also unscientific in evaluating work on some basis other than merit. As a result, they risk hurting the openness that is necessary for international relationships to succeed.

To a large extent, these questions are in the eye of the beholder. Primate research is under legal threat in several European countries.

Would critics attack European researchers who travel to the United States to continue their work, in the same way as researchers who travel to a country with fewer regulations on human embryo research?

For better or worse, legal, ethical and even scientific standards are a patchwork and likely to remain so. Many journals try to accommodate this by requiring that an author adheres to his or her own local laws and ethical standards. IRBs set specific protocols for human and animal research, for instance, on a case-by-case basis, in accordance with variables such as state regulations in the United States.

This patchwork demands international dialogue and an openness to achieving understanding and reaching common ground — none of which is helped by assumptions and veiled prejudice. Realizing what prejudices exist — conscious or not — and then considering whether they are valid for the work in question is a necessary step towards the fair evaluation of science. ■

## Write on

*Biologists are using more informal language in their papers.*

We are not supposed to use first-person pronouns, and contractions aren’t allowed. These rules also discourage unattended anaphoric pronouns and say that split infinitives should be rarely used. And to start a sentence with an initial conjunction is as bad as to include a listing expression, and so on. Exclamation marks are forbidden!

The rules of academic writing are many, but they have one intention: to avoid informal language, in all its forms. Blogs and social media may encourage authors to write it as they say it, but much of what passes for scholarly and scientific prose is simply not designed for human ears. Academic writing is code, with freedom of expression and emotional range curtailed in favour of explicit meaning and a necessary lack of ambiguity. If nothing else, it (by which we mean academic writing, for those still on the watch for unattended pronouns) is writing that knows its audience and gives them what it (the audience) expects.

But, to use a direct question, another stylistic tool on the banned list, is this academic supply and demand still in place? Do the academics of the Internet age still communicate as stiffly as their colleagues did at the time of the Apollo programme? Or, heaven forbid, has some scruffy informality crept into scholarly discourse?

Yes, and no, according to an illuminating new analysis. Formal language is largely intact, the study finds, give or take a mildly more tolerant attitude to split infinitives and initial conjunctions. Yet there has been an explosion in the use of the first-person pronouns in academic

papers by biologists. What, we wondered, is that all about?

The analysis, published online in the journal *English for Specific Purposes*, looked at the language of academic papers selected at random from several high-impact journals published across a range of disciplines in 1965, 1985 and 2015 (K. Hyland and F. Jiang *Engl. Specif. Purp.* <http://doi.org/bssn>; 2017). If anything, academic publishing in applied linguistics and sociology has become slightly more formal. The number of informal features included in papers in the major electrical-engineering journals went up by 9% over the 40 years. But it was the eye-catching increase of 24% in biology journals that stood out, dominated by a headline threefold increase in words such as I and we.

Personal pronouns are frowned on in academic text, with many guidelines to help novice writers avoid them, chief of which is the use of the passive voice (so we did not see something — instead, it was seen). One explanation for the rise is that as the passive voice becomes less fashionable, one obvious way to restructure sentences is to reach for a personal pronoun.

The passive voice is encouraged in scholarly prose precisely for the reasons that dramatists and journalists try to avoid it: it introduces distance between the action and the protagonist and between writer and reader. This, convention suggests, lends an air of detached objectivity to observations and conclusions. It, perhaps, just feels more scientific. The increased use of I and we, the study authors suggest, could also reflect wider language changes in society, or perhaps is down to the increased number of articles written by people for whom English is not their first language. They may not feel so acutely the sense that writing I or we makes a statement of projected authority.

Another explanation is more subtle. Perhaps modern biologists, under increased pressure and competition, do not feel confident that merely stating their case is enough. Personal language builds a connection to the reader and helps, ultimately, to persuade. We think so. Don’t you? ■





## The insect crisis we can't ignore

*We must start an ambitious and professional global programme to explore and preserve invertebrate biodiversity, says Axel Hochkirch.*

According to the International Union for Conservation of Nature (IUCN) Red List of Threatened Species, which I help to compile, just 394 insect species are classed as extinct. That figure is ludicrously low — it's a good bet that several dozen have disappeared in the past week or so alone.

Nobody knows exactly how many invertebrates are vanishing from the planet, but estimates range from 1 to 100 per day. These alarming numbers don't touch our hearts. It is comparable to noticing the number of deaths in a war without seeing the faces of the fallen. So which species are we losing? Is the extinction rate going up or down? What could the implications be? The answer to each question is just as worrying: we have no idea.

We need to preserve invertebrates, not only because they provide valuable ecosystem services such as pollination, pest control and nutrient cycling, but also because every single species on Earth has a right to exist. Many invertebrate species have amazing life histories and are just as charismatic as larger animals. Unknown to many, for example, many grasshoppers woo mates with elaborate courtship songs and dances. It's beautiful!

Targets to assess and protect species already exist. The IUCN has set a 'barometer of life' goal to assess the conservation status of 45,000 invertebrate species by 2020 (S. N. Stuart *et al. Science* **328**, 177; 2010). And more broadly, the parties to the Convention on Biological Diversity have agreed to try to halt biodiversity loss by 2020; progress will be discussed at a meeting next month.

Neither will happen if invertebrate conservation continues to rely on volunteers to assess the status of most insects. This is a serious issue and it demands serious, professional attention.

Governments, science funders and environmental agencies need to invest in one (or, better, several) large centre for invertebrate conservation that employs many proper entomologists to describe species, study their distribution, ecology and threats, and assess their conservation status. A major objective would be liaising with people on the ground to encourage and guide action, build capacity and raise awareness of the global and national importance of invertebrates.

Such a centre would probably cost about €10 million (US\$11 million) to €15 million to establish, and €20 million per year to support. This is comparable to the budgets of similar institutions working at the science-policy interface, such as Germany's Potsdam Institute for Climate Impact Research (which has a yearly budget of about €22 million). Conservation of global biodiversity should be worth such amounts, and is necessary to allow countries to reach their global conservation targets.

Nations are willing to invest in conservation: witness the launch in 2014 of a €20-million tiger-conservation project supported by the

German state-owned bank KfW. Although the status of the tiger is still of concern, strong efforts in conservation will probably help to preserve it. Yet many tiger beetles may go extinct unnoticed.

Conservationists also need to explore and mobilize support from companies that depend on invertebrates as sources of pollination and chemical compounds, and even as inspiration for brands and logos.

Invertebrate extinctions go unnoticed because the number of experts who work on most insect groups is extremely low: it has even reached zero in many regions and for some insect groups. So although the extinction risk of about 18,000 invertebrate species has so far been assessed for the IUCN Red List, nearly one-third are classified as 'data deficient'. Many invertebrate species have not been documented since their original description, and we have no idea whether they still exist.

These large knowledge gaps are worrisome because we can only preserve what we know. Humanity spends a lot of money searching for life in space, but we have not described even one-tenth of the species on Earth.

It is often argued that the lack of invertebrate conservation is caused by invertebrates' low profile compared with charismatic large vertebrates, but this is only part of the truth. Many park managers, rangers and naturalists are very interested in invertebrates, but lack the knowledge and resources to investigate and protect them.

Experience shows that conservation action for invertebrates often starts as soon as information on their identity and threats becomes available. The publication of field guides in central and northern Europe since the 1980s has led to rapid data accumulation, with subsequent publication of national red lists and the establishment of conservation projects. We need to spread these steps

to the species-rich tropical, subtropical and Mediterranean regions. Conservation action to protect invertebrates shows rapid results, due to the species' short generations and small geographical ranges.

One reason for the apparently low number of documented invertebrate extinctions is that scientists are reluctant to declare them. Non-existence cannot be proved, and a few species thought to be extinct have been rediscovered. Prematurely declaring a species extinct may end all conservation efforts and lead to the 'Romeo error', whereby we give up on a species that actually still exists. That is a concern for invertebrates, but there is another Romeo error: some partings are not sweet, just sorrow. We should remember that. ■

---

**Axel Hochkirch** is chair of the IUCN SSC Invertebrate Conservation Subcommittee and co-chair of the IUCN SSC Grasshopper Specialist Group. He works at Trier University in Germany.  
e-mail: hochkirch@uni-trier.de

HUMANITY SPENDS  
A LOT OF MONEY  
SEARCHING FOR  
**LIFE IN SPACE,**  
BUT WE HAVE NOT  
DESCRIBED  
ONE-TENTH OF THE  
**SPECIES ON  
EARTH.**



# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## METABOLISM

### Fat-molecule drop boosts metabolism

Metabolic health can be improved in mice by inhibiting the formation of compounds called ceramides in fat cells under the skin.

Ceramides are fatty molecules that have been associated with obesity and metabolic disease. Bhagirath Chaurasia at the University of Utah in Salt Lake City and his colleagues found that obese mice that could not make ceramides in fat tissue had improved energy metabolism. The mice also showed decreased inflammation and increased sensitivity to insulin (diminished sensitivity is a hallmark of type 2 diabetes). The reduction of ceramide levels in subcutaneous fat cells was linked to a cellular shift from an energy-storage mode to an energy-burning one.

The authors also report that patients with obesity and type 2 diabetes had more ceramides in subcutaneous fat cells than people of the same weight without diabetes.

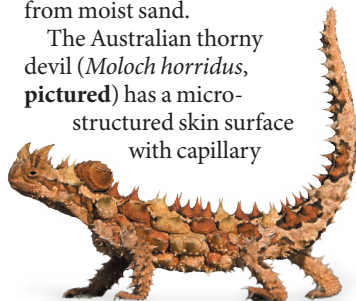
*Cell Metab.* <http://doi.org/bsn2> (2016)

## ZOOLOGY

### Thorny devil may drink from sand

A desert lizard survives in arid conditions by harvesting and drinking water from puddles, and possibly even from moist sand.

The Australian thorny devil (*Moloch horridus*, pictured) has a micro-structured skin surface with capillary



channels that collect and transport water to the animal's mouth. Philipp Comanns at RWTH Aachen University, Germany, and his colleagues found that this capillary system fills up completely when the animal stands in a puddle, allowing the lizard to drink the water. Conversely, when the lizard stands on moist sand, the capillary network fills to only about 59% of its capacity, and the animal is not able to drink this water. But tests of wet sand on skin replicas showed that moisture can be extracted from sand and pulled into the

capillaries by gravity to fill them.

*M. horridus* has been observed in the wild shovelling wet sand onto its back, and the authors say this increases the contact area between skin and sand, boosting the volume of water it can collect and drink. *J. Exp. Biol.* 219, 3473–3479

## EVOLUTION

### Cat DNA shaped by diet

Carnivores have experienced stronger natural selection than plant-eating animals, perhaps

measured the way the seismic waves travelled through the ground. They confirmed that the rock beneath the mountain is too cold to create magma. Instead, the molten rock that feeds the volcano seems to come from off to the east.

The sideways plumbing helps to explain why the mountain remains active despite lying to the west of most neighbouring volcanoes.

*Nature Commun.* 7, 13242 (2016)



## GEOPHYSICS

### Volcano sneaks in magma sideways

The Mount St Helens volcano in the United States has crooked plumbing. Rather than sitting above the magma source that feeds it, the mountain is off to one side.

A major eruption in 1980 saw 57 people killed by the volcano, which is in Washington state. Steven Hansen at the University of New Mexico in Albuquerque and his colleagues set off explosives around Mount St Helens and

because of their limited diet.

Joo-Hong Yeo at the National Institute of Biological Resources in the Republic of Korea and his colleagues compared the genomes of 18 mammals, including carnivores such as leopards, omnivores such as humans, and herbivores such as giant pandas. They found that carnivore genomes have lost many genes for carbohydrate digestion, and share changes to genes involved in muscle strength and agility, making them good hunters. By contrast, omnivore and herbivore genomes shared

SUNSET AVENUE PRODUCTIONS/GETTY

YVA MOMATIUK & JOHN EASTCOTT/  
MINDEN PICTURES/FLPA



fewer adaptations with others in their groups, suggesting that their diets have imposed weaker selection compared with that of carnivores.

Many big cats showed recent losses in genetic diversity, suggesting that population declines may be linked to their strict diet.

*Genome Biol.* 17, 211 (2016)

## ASTROPARTICLE PHYSICS

## Cosmic rays make more muons

Particle showers caused by natural ultra-high-energy collisions in Earth's atmosphere produce more muons — heavier cousins of the electron — than current physics models can explain.

Using the Pierre Auger Observatory in Argentina, Glennys Farrar of New York University and her colleagues studied showers of particles produced when 411 ultra-high-energy cosmic rays — atomic nuclei thought to originate outside the Galaxy — collided with air molecules. They also studied the fluorescent light the cascades created. The team found that the collisions, which are 10 times more energetic than those generated at the Large Hadron Collider (LHC) near Geneva in Switzerland, produced 30–60% more muons than simulations based on LHC results predict.

The results suggest that either the underlying models contain flaws, or that physics is fundamentally different at these higher energies.

*Phys. Rev. Lett.* 117, 192001 (2016)

## PLANT BIOLOGY

## Plants transmit light down stems

Plant stems can act like fibre-optic cables, piping light from above ground down to the roots.

Plants' roots produce proteins that respond to light, but it was not clear how light reached below the ground.

Chung-Mo Park of Seoul National University, Ian Baldwin of the Max Planck Institute for Chemical Ecology in Jena, Germany, and their team investigated this in thale cress (*Arabidopsis thaliana*). They found that illuminating the plant shoot altered gene expression in the roots, even when they prevented light from shining through the soil. Light in the red to near-infrared range was efficiently conducted through stem and root tissues.

Plants bearing a mutation in a light-responsive protein called HY5 showed abnormal root growth in response to shoot illumination, suggesting that light-sensing in plants is necessary for normal root development.

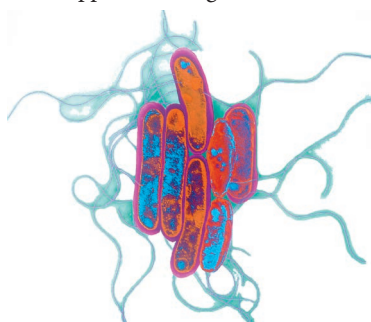
*Sci. Signal.* 9, ra106 (2016)

## MICROBIAL EVOLUTION

## Legionnaires' strains adapt well

Bacteria responsible for many cases of Legionnaires' disease emerged only in recent decades and seem to be adapting to human environments.

*Legionella pneumophila* (pictured) is found in aquatic environments worldwide and can contaminate water supplies, causing outbreaks of



pneumonia that can be fatal. A team led by Julian Parkhill at the Wellcome Trust Sanger Institute in Hinxton, UK, and Carmen Buchrieser at the Pasteur Institute in Paris sequenced the genomes of 337 *L. pneumophila* isolates belonging to 5 types that cause almost half of all cases of Legionnaires' disease in northwest Europe. Sequence



analysis suggested that the 5 types emerged independently during the past few decades and spread around the world.

The recent emergence and spread of these lineages suggests that people infected with the bacterium are helping to disseminate it and that it is adapting to man-made water systems, the authors say.

*Genome Res.* 26, 1555–1564 (2016)

## ZOOLOGY

## Old apes can't see up close

The discovery that bonobos lose their close vision with age, as humans do, offers clues to the cause of this visual decline.

Heungjin Ryu of Kyoto University in Japan and his colleagues took photos of 14 wild bonobos ranging between 11 and 45 years of age as they groomed family and friends (pictured) — a behaviour that requires careful inspection. The team found that the distance from eyes to fingers was significantly longer in older animals than in younger ones. Moreover, a comparison with older images of grooming for one of the individuals, called Ki, showed that the distance had increased from about 12 centimetres in 2009 to roughly 17 centimetres in 2015. Bonobos tended to exhibit this 'long-sighted' condition, known as presbyopia, from about age 40.

The results suggest that presbyopia is caused not by

human activities such as reading and computer work, but by natural ageing of the primate eye.

*Curr. Biol.* 26, R1–R3 (2016)

## FLUID DYNAMICS

## Fluid spills from strange straws

A straw's shape has a counter-intuitive effect on the ability of fluids to flow inside it, according to a theoretical study.

If a cylindrical container with open ends is turned on its side, a liquid contained in it tends to stay inside because of surface tension, provided the container is narrow enough. However, Carlos Rascón of Carlos III University in Madrid and his collaborators have shown that fluid in containers with certain other shapes — those that are elliptical or triangular in cross-section, for example — will flow out no matter how narrow the tubes are. The team calculated that surface-adhesion forces will work together with gravity to make fluids spread along the entire length of the vessel, and ultimately escape.

The researchers say the finding has potential applications in microfluidic devices, and plan to test their predictions experimentally.

*Proc. Natl Acad. Sci. USA*

<http://doi.org/bspc> (2016)

## NATURE.COM

For the latest research published by Nature visit:

[www.nature.com/latestresearch](http://www.nature.com/latestresearch)

# SEVEN DAYS

The news in brief

## EVENTS

### Record smasher

The world's most powerful particle collider has set a new record and surpassed its own target for data collection. CERN, Europe's particle-physics lab near Geneva, Switzerland, announced on 1 November that the Large Hadron Collider (LHC) performed around 7 quadrillion ( $10^{15}$ ) proton–proton collisions during its 2016 run, which ended on 25 October. This is 60% higher than its target for the period, and the amount of data collected by the detectors in 2016 exceeded the total for 2010 to 2015. The LHC will now smash together protons and lead ions for a month, before shutting down until May 2017 for repairs and upgrades.

### Korean ship launch

South Korea launched its huge new research vessel *Isabu* on 2 November. The 5,900-tonne ship is substantially larger than the country's previous flagship, the 1,422-tonne *Onnuri*. During *Isabu*'s maiden voyage to the northwest Pacific Ocean and the Indian Ocean, researchers aboard will look at the effects of climate change on currents and study organisms living near geothermal springs. The vessel, which is equipped to perform sea-floor-penetrating seismic surveys and to collect sediment cores up to 30 metres long, heralds an expansion of the country's marine-science capabilities.

## SPACE

### Heavy lifter

China launched its biggest rocket yet on 3 November, taking it a step closer to assembling a permanently

crewed space station and sending large probes on interplanetary missions. The Long March 5 lifted off from the Wenchang launch site on Hainan Island, carrying a spacecraft to test electric propulsion systems. The launcher's ten engines provide nearly the same lift capacity as that of the most powerful rocket currently in use, the US-built Delta IV Heavy. Long March 5 is expected to carry a Chinese space station into orbit in stages starting in 2018.



SAJJAD HUSSAIN/AFP/GETTY

## Delhi choked by worst smog in years

Delhi has endured several days of heavy smog — reportedly the city's worst in 17 years. Measurements at the US embassy in the Indian capital showed an air-quality-index value of 999; around 300 is considered hazardous. Levels of harmful particles known as PM<sub>2.5</sub> (those with a diameter of 2.5 micrometres or less) reached as high as 762 micrograms per cubic metre in a part of the city. The World Health Organization gives 25 micrograms per cubic metre as a guideline average limit. Among the causes blamed have been

construction dust, rising traffic, the burning of agricultural residues in neighbouring states, and the Hindu festival of Diwali, during which firework celebrations are customary. Delhi's government announced several temporary emergency measures, including the closure of schools for three days, a halt on construction and restrictions on the numbers of vehicles on roads. Delhi is one of the world's most polluted cities, according to the WHO. Estimates suggest that particulate air pollution causes one-tenth of the city's deaths.

## POLICY

### UK pollution case

A High Court judge has ruled that the UK government is not doing enough to combat air pollution, and is relying on over-optimistic pollution modelling. ClientEarth, a group of environmental law activists, argued that the environment minister had failed to take steps to address levels of nitrogen dioxide "as soon as possible", as required by European Union law. Nitrogen dioxide is largely

emitted by road vehicles. The ruling, handed down on 2 November, is the second defeat on the issue for the government in 18 months.

### Plant patents

The European Commission has clarified its legal position on whether plant strains created using classical breeding and selection methods are patentable. They are not, it announced on 3 November. That puts the European Union at odds with the European Patent Office



ALLEN J. SCHABEN/LA TIMES/GETTY

(EPO), which after a long battle decided in March last year that they were. All EU member states belong to the EPO, which also has a further ten members. The office said that it would be considering how to align the positions. Around 320 applications for plant products created with biological methods (those that do not involve genetic engineering) are in the EPO pipeline.

## Climate pact begins

The landmark Paris climate agreement, adopted in the French capital in December 2015 after years of complex international negotiations, officially took effect on 4 November. One hundred nations, accounting for 69% of global greenhouse-gas emissions, have now formally joined the deal, which aims to limit global warming to less than 2 °C above pre-industrial levels. The first meeting of the 197 parties to the Paris agreement is scheduled to take place at the United Nations Climate Change Conference in Marrakesh, Morocco, which runs until 18 November.

### PEOPLE

## Former NAS chief

Noted atmospheric chemist Ralph Cicerone, who served as president of the US National Academy of Sciences from



July 2005 until June this year, died on 5 November. Cicerone (pictured) was a leading advocate for educating the public and politicians about topical issues such as evolution and global warming. He led a set of influential reports laying out the threats and possible responses to climate change. In the 1970s, Cicerone helped to show that chlorine atoms can break apart stratospheric ozone — a discovery that paved the way for banning chlorofluorocarbon compounds to protect the ozone layer.

### RESEARCH

## Trial tracker

A study that used an automated tool to trawl through thousands of records in the world's leading clinical-trials database has revealed which drug firms and academic institutions are failing to publish the results of trials. Although

the failure is already well documented, emerging software can perform a more comprehensive search than was previously possible. Ben Goldacre and Anna Powell-Smith at the University of Oxford, UK, developed the tool to search the ClinicalTrials.gov database for trials that were completed at least two years ago. The trawl found that, of 26,000 trials evaluated, 45.2% had no published results. The work, which has not yet been peer reviewed, was published on 3 November (A. Powell-Smith and B. Goldacre *F1000Research* <http://doi.org/bsnn>; 2016). See page 153 for more.

## Illegal ivory

The illegal ivory trade depends almost entirely on elephants that have been recently killed, according to researchers who have carbon-dated hundreds of seized ivory tusks. Some had wondered whether corrupt governments were contributing to the trade by selling off old ivory, bit by bit, from stockpiles. Thure Cerling at the University of Utah in Salt Lake City and his colleagues measured the decay of carbon-14 isotopes in 231 ivory tusks, confiscated between 2002 and 2014, to determine when the elephants they were taken from had died. Their analysis, published on

## COMING UP

### 17 NOVEMBER

Astronauts Peggy Whitson and Thomas Pesquet and cosmonaut Oleg Novitskiy head to the International Space Station aboard a Russian Soyuz craft, launching from the Baikonur Cosmodrome in Kazakhstan.

### 22–24 NOVEMBER

Researchers from across disciplines discuss the scientific and clinical aspects of sleep disorders at the 2016 Sleep Summit in London.

[go.nature.com/2fa0xoz](http://go.nature.com/2fa0xoz)

8 November, found that only four specimens were more than five years old at the time they were seized (T. E. Cerling *et al. Proc. Natl Acad. Sci. USA* <http://doi.org/bssc>; 2016).

### FUNDING

## Social climbing

The Singapore government plans to invest 350 million Singapore dollars (US\$252 million) over five years in its recently formed Social Science Research Council, deputy prime minister Tharman Shanmugaratnam announced on 7 November. The council, formed last year, was launched to boost social sciences and humanities research into challenges including social integration and spurring productivity, as well as to help build up local research talent. The funding from Singapore's Ministry of Education represents a 45% increase in the ministry's spending on these fields compared with the previous five years, said Shanmugaratnam, who is also coordinating minister for economic and social policies.

**NATURE.COM**

For daily news updates see:

[www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

An online poll answered by thousands of *Nature* readers shows that almost two-thirds have considered quitting research, and that 15% have actually quit. The poll accompanied a News Feature published on 26 October (*Nature* 538, 446–449; 2016). The fight for funding was the greatest challenge, readers said. Almost 40% of respondents said that they work more than 60 hours a week. Many also felt that they lacked a work–life balance and that progression is judged too heavily on publication record.

## SHOULD I QUIT?

Almost two-thirds of *Nature*'s readers say they have considered quitting research; 15% have actually quit.



SOURCE: NATURE

# NEWS IN FOCUS

**UK POLICY** Scientists anxiously await Brexit chancellor's speech **p.148**

**BIBLIOMETRICS** Little-known algorithm gains influential following **p.150**

**ECOLOGY** Quest to record the African soil microbiome begins **p.152**



**DEMENTIA** How to avoid a health-system meltdown **p.156**

JOHN STANMEYER/NGC



The increasing rate of deforestation in the Brazilian Amazon has driven up the country's greenhouse-gas emissions.

POLICY

## Brazil debates loosening environmental protections

*Barrage of proposals would allow developers to sidestep environmental reviews.*

BY JEFF TOLLEFSON

Environmentalists in Brazil are feeling the heat. Conservative lawmakers want to weaken the country's environmental regulations to clear the way for rapid development of energy facilities, mines and agriculture — in the Amazon and beyond. Their push comes at a time of economic and political turmoil following the impeachment in August of former President Dilma Rousseff.

"It's an offensive against our regulatory system," says Mauricio Guetta, an attorney with the Socio-Environmental Institute, an advocacy group in São Paulo.

More than 20 legislative proposals are circulating in the Brazilian Congress to loosen regulations governing activities such as building roads and hydroelectric dams or expanding agricultural businesses. One proposed constitutional amendment would ensure approval of a project once its developers

have submitted an environmental-impact analysis — essentially eliminating government review. That proposal has stalled in the Senate, but the government of President Michel Temer is developing its own legislation to overhaul the environmental-licensing system, which many consider ineffective.

"Something will happen, most probably in the wrong direction," says Nilvo Silva, a former head of the licensing division of the Brazilian Institute of Environment and ▶



► Renewable Natural Resources, an environmental-enforcement agency.

The debate comes during Brazil's worst recession in decades, and follows corruption scandals that brought down Rousseff and her leftist Workers' Party. The Brazilian Democratic Movement Party has taken the reins but it, too, has been tainted. Several cabinet members have resigned, and corruption investigations are continuing — with Temer in the crosshairs.

The embattled president has promised to maintain Brazil's environmental agenda, including its commitments under the Paris climate agreement. But agricultural and business interests are pushing back against environmental protections set by the Workers' Party under Rousseff's predecessor, Luiz Inácio Lula da Silva, endangering more than a decade of progress on issues such as deforestation.

"They keep paying lip service to environmental issues, but we can't be confident in the implementation of policies," says Paulo Barreto, a senior researcher at the Amazon Institute of People and the Environment, an activist group in Belém.

Much of the concern centres on the Amazon, where the rate of forest loss has increased by nearly 36% since 2012. More than 6,200 square kilometres of land were cleared for agriculture in 2015, and many expect that

number to increase when the 2016 data are released next week. The deforestation helped to increase Brazil's overall greenhouse-gas emissions by 3.5% in 2015, even as emissions from the energy sector fell, according to a 27 October report by the Climate Observatory, a coalition of advocacy groups in São Paulo.

Brazil's environment minister, José Sarney Filho, says that some people may be taking advantage of the political crisis to clear forest. The government has responded by bolstering funding to enforce existing laws. "We expect that we will once more be on the right track of reducing deforestation," he adds.

Barreto says that part of the problem stems from changes to Brazil's forest law in 2012 that weakened rules and let many landowners off the hook for past violations. The latest efforts to streamline the environmental licensing system would further advance that agenda.

One project that could be fast-tracked if the latest regulatory changes take effect is the proposed Volta Grande mine on the Xingu River in the Amazonian state of Pará, near the controversial Belo Monte hydroelectric dam. The Volta Grande project, which would be Brazil's largest gold mine, is facing legal challenges from independent prosecutors who say that the government's analysis of its social and environmental impacts was flawed.

But it would be difficult to fight such projects in the courts if the proposed constitutional amendment were enacted, says Raffael Tófoli, an ecologist at the State University of Maringá.

Many scientists and environmentalists acknowledge that Brazil's regulatory system is slow and often ineffective. The solution is to improve environmental assessments, increase public participation in environmental reviews and give regulators more resources, says Luis Sánchez, an engineer at the University of São Paulo who conducts environmental assessments. "This is something that could be solved without changing the law," he says.

It's not yet clear what solutions Temer's government will propose. Green groups say that the environment ministry's first draft of a proposal to reform the licensing process expanded the focus from individual projects to the social and environmental effects of development across an entire landscape. But that proposal is now circulating among other ministries that oversee activities such as mining, energy and infrastructure, and some observers say that the latest leaked drafts show that the plan is being watered down with concessions to industry.

"We are all waiting for the government to present this bill," says Guetta, "but we are seeing the text get worse every day." ■

SEE EDITORIAL P.139

## UK POLICY

# Brexit chancellor's annual address is science nail-biter

*Government's first Autumn Statement could reveal how it regards science.*

BY ELIZABETH GIBNEY

Since the United Kingdom voted to leave the European Union, science has existed under a cloud of uncertainty. The future of international collaboration and the mobility of scientists is in limbo. And the government overhaul that followed the June vote included a

reshuffle of government ministers, the creation of two departments to implement Brexit and a new prime minister.

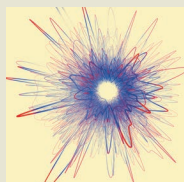
Hints of how the new guard (see 'Science in flux') regards science could come on 23 November, when the Chancellor of the Exchequer Philip Hammond lays out his financial plans in the Autumn Statement. The

annual address on the nation's finances will be the first from the new 'Brexit government'.

"The Autumn Statement is a real nail-biter for scientists," says cell biologist Jennifer Rohn at University College London, an executive board member of the lobby group Science is Vital. "In the past, even when we've been unsure of the specific outcomes, we've at



## NATURE VIDEO



Scientists can publish their best work at any age [go.nature.com/2fjqpaw](http://go.nature.com/2fjqpaw)

## MORE NEWS

- Algorithm quantifies gender bias in astronomy [go.nature.com/2fowmlu](http://go.nature.com/2fowmlu)
- Psychologists argue about whether smiling makes cartoons funnier [go.nature.com/2fjqpwi](http://go.nature.com/2fjqpwi)
- Illegal ivory mostly from recent elephant killings [go.nature.com/2fjowzn](http://go.nature.com/2fjowzn)

## NATURE PODCAST



A brain-spine interface for walking; a CERN for neuroscience; and a climate tax for food [nature.com/nature/podcast](http://nature.com/nature/podcast)

IMAGE: KIM ALBRECHT; DATA IMAGE: ROBERTA SINATRA

► Renewable Natural Resources, an environmental-enforcement agency.

The debate comes during Brazil's worst recession in decades, and follows corruption scandals that brought down Rousseff and her leftist Workers' Party. The Brazilian Democratic Movement Party has taken the reins but it, too, has been tainted. Several cabinet members have resigned, and corruption investigations are continuing — with Temer in the crosshairs.

The embattled president has promised to maintain Brazil's environmental agenda, including its commitments under the Paris climate agreement. But agricultural and business interests are pushing back against environmental protections set by the Workers' Party under Rousseff's predecessor, Luiz Inácio Lula da Silva, endangering more than a decade of progress on issues such as deforestation.

"They keep paying lip service to environmental issues, but we can't be confident in the implementation of policies," says Paulo Barreto, a senior researcher at the Amazon Institute of People and the Environment, an activist group in Belém.

Much of the concern centres on the Amazon, where the rate of forest loss has increased by nearly 36% since 2012. More than 6,200 square kilometres of land were cleared for agriculture in 2015, and many expect that

number to increase when the 2016 data are released next week. The deforestation helped to increase Brazil's overall greenhouse-gas emissions by 3.5% in 2015, even as emissions from the energy sector fell, according to a 27 October report by the Climate Observatory, a coalition of advocacy groups in São Paulo.

Brazil's environment minister, José Sarney Filho, says that some people may be taking advantage of the political crisis to clear forest. The government has responded by bolstering funding to enforce existing laws. "We expect that we will once more be on the right track of reducing deforestation," he adds.

Barreto says that part of the problem stems from changes to Brazil's forest law in 2012 that weakened rules and let many landowners off the hook for past violations. The latest efforts to streamline the environmental licensing system would further advance that agenda.

One project that could be fast-tracked if the latest regulatory changes take effect is the proposed Volta Grande mine on the Xingu River in the Amazonian state of Pará, near the controversial Belo Monte hydroelectric dam. The Volta Grande project, which would be Brazil's largest gold mine, is facing legal challenges from independent prosecutors who say that the government's analysis of its social and environmental impacts was flawed.

But it would be difficult to fight such projects in the courts if the proposed constitutional amendment were enacted, says Raffael Tófoli, an ecologist at the State University of Maringá.

Many scientists and environmentalists acknowledge that Brazil's regulatory system is slow and often ineffective. The solution is to improve environmental assessments, increase public participation in environmental reviews and give regulators more resources, says Luis Sánchez, an engineer at the University of São Paulo who conducts environmental assessments. "This is something that could be solved without changing the law," he says.

It's not yet clear what solutions Temer's government will propose. Green groups say that the environment ministry's first draft of a proposal to reform the licensing process expanded the focus from individual projects to the social and environmental effects of development across an entire landscape. But that proposal is now circulating among other ministries that oversee activities such as mining, energy and infrastructure, and some observers say that the latest leaked drafts show that the plan is being watered down with concessions to industry.

"We are all waiting for the government to present this bill," says Guetta, "but we are seeing the text get worse every day." ■

SEE EDITORIAL P.139

## UK POLICY

# Brexit chancellor's annual address is science nail-biter

*Government's first Autumn Statement could reveal how it regards science.*

BY ELIZABETH GIBNEY

Since the United Kingdom voted to leave the European Union, science has existed under a cloud of uncertainty. The future of international collaboration and the mobility of scientists is in limbo. And the government overhaul that followed the June vote included a

reshuffle of government ministers, the creation of two departments to implement Brexit and a new prime minister.

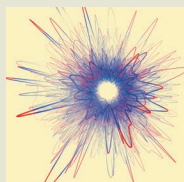
Hints of how the new guard (see 'Science in flux') regards science could come on 23 November, when the Chancellor of the Exchequer Philip Hammond lays out his financial plans in the Autumn Statement. The

annual address on the nation's finances will be the first from the new 'Brexit government'.

"The Autumn Statement is a real nail-biter for scientists," says cell biologist Jennifer Rohn at University College London, an executive board member of the lobby group Science is Vital. "In the past, even when we've been unsure of the specific outcomes, we've at



## NATURE VIDEO



Scientists can publish their best work at any age [go.nature.com/2fjqpaw](http://go.nature.com/2fjqpaw)

## MORE NEWS

- Algorithm quantifies gender bias in astronomy [go.nature.com/2fowmlu](http://go.nature.com/2fowmlu)
- Psychologists argue about whether smiling makes cartoons funnier [go.nature.com/2fjqpwi](http://go.nature.com/2fjqpwi)
- Illegal ivory mostly from recent elephant killings [go.nature.com/2fjowzn](http://go.nature.com/2fjowzn)

## NATURE PODCAST



A brain-spine interface for walking; a CERN for neuroscience; and a climate tax for food [nature.com/nature/podcast](http://nature.com/nature/podcast)

IMAGE: KIM ALBRECHT; DATA IMAGE: ROBERTA SINATRA





CHRISTOPHER GOODNEY/BLOOMBERG VIA GETTY

UK Chancellor of the Exchequer Philip Hammond will make his first financial statement on 23 November.

least been familiar with the Treasury's general stance on the importance of science. With a change of guard, everything is up in the air."

"The signalling of intent is almost more important than sums of money that may or may not be dished out," adds Sarah Main, director of the Campaign for Science and Engineering in London. "We'll be looking for political signals and financial commitments." Main is even "quite hopeful" that a new agenda might provide fresh science opportunities.

Former chancellor George Osborne was vocal about the importance of blue-sky research, and protected the science budget from cuts that hit many areas under his austerity programme. But he did not increase investment in real terms and during his tenure the United Kingdom dropped to the bottom of the G8 countries in terms of the percentage of gross domestic product its government spends on research and development (R&D).

Hammond may be more willing than Osborne to spend public money, having already announced that he does not plan to hit Osborne's targets for cutting the United Kingdom's deficit. Whether this will translate into increased funds for science is unclear. Stephen Curry, a structural biologist at Imperial College London and a member of the advisory board for Science is Vital, describes Hammond — who wanted to remain in the EU — as a "fairly steady head". But he says the jury is out on whether he will be a science-friendly chancellor.

Other positive signals come from newly appointed business secretary and former science minister Greg Clark, who told Parliament on 31 October that being at "the cutting edge of research and development" was a cornerstone of the new government's industrial strategy, a range of policies

aimed at boosting the economy.

Main sees the strategy as an opportunity for the government to increase the country's investment in R&D. "At the moment there's a stalemate with both government and industry saying the other needs to invest more," she says. But she cautions that the industrial strategy could also tempt the government to predict, and then pick, specific 'winning technologies' to fund, instead of boosting the research base more broadly.

Others doubt that science will be regarded as important enough to feature in the Autumn Statement. Navigating the path to Brexit is "more than a full-time job," says Kieron Flanagan, a science-policy researcher at the University of Manchester. He notes that after the Brexit vote, it took the government seven weeks

to announce that it would guarantee replacement funding for existing EU-funded research projects, even after the United Kingdom leaves the EU. "That it took so long to agree on that shows they're elsewhere," he says. If science is mentioned in the statement, he predicts it is likely to be "an easy, shiny project to announce" rather than anything of substance.

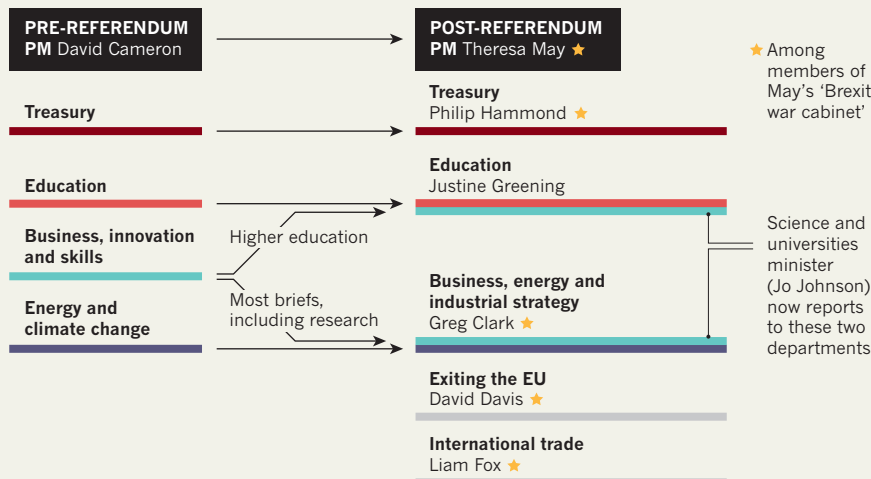
Funding is only one direction in which UK science needs reassurance. "It's just as important for the government to start making friendly noises to our overseas colleagues in a way that they distinctly haven't been over the past month or two," says Curry. The terms on which the United Kingdom leaves the EU and builds new international relations — to be negotiated by two new bodies, the Department for Exiting the European Union and the Department for International Trade — will dictate whether the country remains part of EU research programmes post-Brexit.

Scientists should also be watching the Brexit 'war cabinet', a collection of government ministers that Prime Minister Theresa May has tasked with making the ultimate decisions on the EU exit, according to leaked documents. Although Hammond, Clarke and Damian Green — who made the case for why remaining would be good for universities ahead of the referendum — are members, they might find themselves sidelined by the half of the committee who campaigned to leave the EU, says Mike Galsworthy, who is programme director for the lobby group Scientists for EU in London. "If there is war within the 'war' cabinet, then the champions of universities, European collaboration and science investment will lose."

Main notes that the changing of the guard requires organizations like hers to set out the value of science to the economy and to society all over again. "There is a whole new cast of characters now for us to meet with and make that case to." ■

## SCIENCE IN FLUX

The United Kingdom's vote to leave the European Union triggered a government shake-up. Here are some of the departmental changes and newly appointed ministers that matter to science.



## BIBLIOMETRICS

# The quiet rise of the NIH's hot new metric

*Biomedical funders worldwide are adopting the US agency's free Relative Citation Ratio to analyse grant outcomes.*

BY GAUTAM NAIK

A little-known algorithm that scores the influence of research articles has become an important grant-management tool at the world's largest biomedical funding agency, the US National Institutes of Health (NIH).

In 2015, the NIH's Office for Portfolio Analysis (OPA) in Bethesda, Maryland, devised the tool to compare the performance of articles from different fields more fairly. Now, one of the NIH's biggest institutes is using the metric — the Relative Citation Ratio, or RCR — to identify whether some types of grant deliver more bang for their buck. Other funders have adopted the RCR, which the agency offers freely online. In the United Kingdom, biomedical charity the Wellcome Trust is using the RCR to analyse its grant outcomes; in Italy, Fondazione Telethon, a charity that supports research into genetic diseases, is testing the RCR as a way to evaluate its funding initiatives.

"It's getting a very good reception both inside and outside the NIH," says George Santangelo, director of the OPA. Santangelo's team, an 18-strong group of scientists, statisticians and data wranglers, was founded five years ago to devise tools to analyse NIH funding opportunities.

Asked to measure which NIH research has had the most influence, the team chose not to judge articles simply by the journal in which

they were published. That approach gives articles in highly cited journals higher scores, but it has acknowledged flaws. An important study might be underestimated because it was not published in an elite journal, for instance. Simply counting citations, meanwhile, fails to capture the idea that articles should be judged relative to similar papers: an algebra paper with a few dozen citations, for example, may have a greater impact in mathematics than a widely cited cancer study would have in oncology.

Algorithms that compare articles with others in their field are offered by commercial analysis firms such as Elsevier, but Santangelo's team argue that its metric is technically as good, if not superior — and, importantly, more accessible. (The NIH has posted help files and its full code online.) "No other metric we've seen is as transparent as RCR," Santangelo says.

The algorithm is complex. It defines an article's research 'field' as the cluster of papers that it has been co-cited with: a dynamic cohort that grows all the time. It then calculates the background field's citation rate — the average citations of the field's journals each year. After a few months of accrued citations, an article's actual performance can then be benchmarked against this background — although in some cases one has to wait a year, says Santangelo.

To put this benchmarking in context, the team compares it to how NIH-funded papers in the same field and year performed (B. I. Hutchins *et al.* *PLOS Biol.* **14**, e1002541;

2016). This boils everything down to a simple number, the RCR. An RCR of 1.0 means that an article has had exactly as much influence as the median NIH-funded paper in the same year and field; 2.0 means a paper has had twice as much influence, and so on (see 'A measure of influence'). Anyone can upload PubMed papers at a website called iCite to find out their RCR score (<https://icite.od.nih.gov>).

The new metric has critics. "Our analysis shows that it is not better than other indicators," says Lutz Bornmann, a bibliometric specialist at Germany's Max Planck Society in Munich. The society has been using at least three other field-normalized metrics for several years to evaluate its institutions, but has no plans to adopt the RCR. It says that the metric is too complicated and too restrictive because it has been applied only to the PubMed database, which contains largely biomedical papers, so doesn't work for physical-sciences analysis.

The RCR, however, is starting to gain ground as an analysis tool. At the US National Institute of General Medical Sciences (NIGMS) in Bethesda, a team used the metric to compare the impact of large, multimillion-dollar 'programme project' grants — which fund teams of researchers — with smaller grants for individual principal researchers. Papers produced by both grants had similar scores. "It has helped us take a very hard look at our support for team science," says NIGMS director Jon Lorsch.

Another question that the NIGMS asked was whether scientists who get more money produce better outcomes than those who get less funding. Again, when the RCR numbers were tallied, it turned out that more NIH money didn't lead to higher-RCR papers. "So maybe we shouldn't fund scientists who are already well-funded," says Michael Lauer, deputy director for extramural research at the NIH.

The tool is also catching on outside the NIH. Jonathon Kram, a research analyst at the Wellcome Trust, says that his group uses the RCR to analyse grants, and to benchmark the performance of the trust's funding schemes against other funders' grants. Unlike other normalized metrics, he says, the RCR "has a transparent methodology and is available free".

Software firm ÜberResearch in Cologne, Germany, has built a database of grants awarded by some 200 funders, and has begun publishing RCR scores for each publication in its grant database. "We use RCR to better judge the history of the researchers listed in our database," says Stephen Leicht, a co-founder of ÜberResearch. (It is part-owned by Digital Science, a firm operated by the Holtzbrinck Publishing Group, which has a share in *Nature's* publisher.)

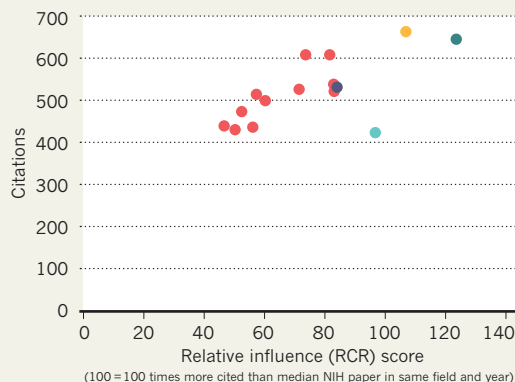
And Fondazione Telethon, the Italian charity, says that it is testing the RCR and hopes to adopt it. "We are not going to use it to help decide funding decisions but more as a tool for analysis," says Lucia Monaco, the charity's chief scientific officer. "We want to make sure that every euro we invest is in excellent research." ■

## A MEASURE OF INFLUENCE

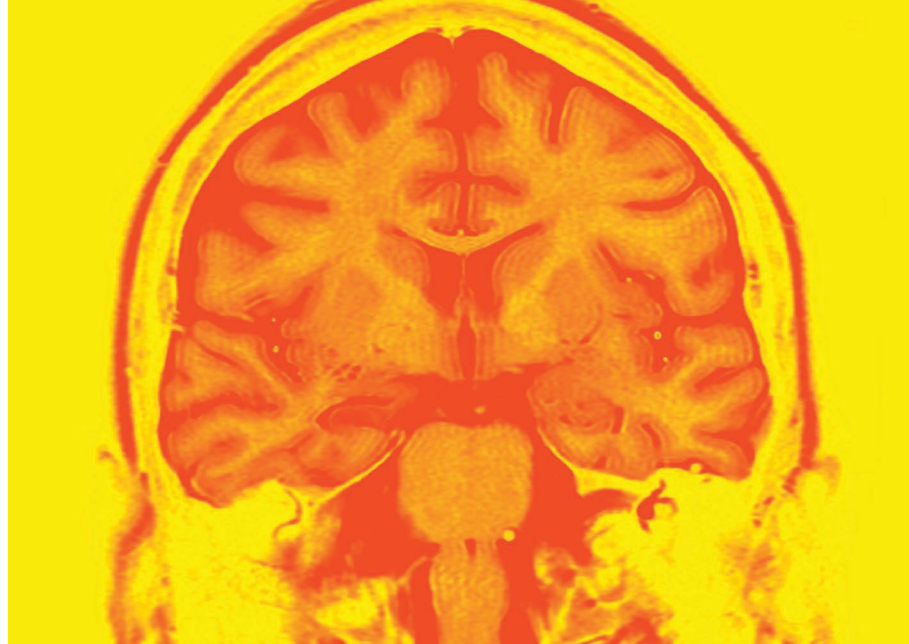
The 15 most-cited papers funded by the US National Institutes of Health in 2014 also score highly on the agency's Relative Citation Ratio (RCR) for measuring influence. But the two are not exactly correlated.

### Highest RCR scores

- 'Diet rapidly and reproducibly alters the human gut microbiome'  
*Nature* **505**, 559–563 (2014).
- 'Biological insights from 108 schizophrenia-associated genetic loci'  
*Nature* **511**, 421–427 (2014).
- '2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation'  
*J. Am. Coll. Cardiol.* **64**, e1–e76 (2014).
- 'ProteomeXchange provides globally coordinated proteomics data submission and dissemination'  
*Nature Biotech.* **32**, 223–226 (2014).







Magnetic resonance imaging can be used to study the brain and diagnose brain disorders.

## NEUROSCIENCE

# Public-health goal for brain mappers

*World Health Organization asks that findings translate into clinical benefits.*

BY SARA REARDON

Major brain-mapping projects have multiplied in recent years, as neuroscientists develop new technologies to decipher how the brain works. These initiatives focus on understanding the brain, but the World Health Organization (WHO) wants to ensure that they work to translate their early discoveries and technological advances into tests and treatments for brain disorders.

“We think there are side branches from projects that could be pursued with a very small investment to benefit public health,” says Shekhar Saxena, director of the WHO’s mental-health and substance-abuse department.

Saxena will make that case on 12 November at the annual meeting of the Society for Neuroscience in San Diego, California — continuing a discussion that began in July at the WHO’s headquarters in Geneva, Switzerland. Among the roughly 70 people who attended that first meeting were leaders of the major brain initiatives, including the US BRAIN (Brain Research through Advancing Innovative Neurotechnologies) Initiative, launched in 2013; the European Human Brain Project, started in 2013; and the Japanese Brain/MINDS project, launched in 2014.

All of these projects focus on basic research on the brain or the development of

sophisticated tools to study it. Clinical applications are an ultimate, rather than an immediate, goal. But at the Geneva meeting, project leaders agreed, in principle, that they should do more to adapt brain-imaging technologies for use in clinical diagnoses.

“The WHO is concerned that the emphasis on building these very expensive devices could worsen the health disparities that we have now between the developed and under-developed world,” says Walter Koroshetz, director of the US National Institute of Neurological Disorders and Stroke, which is part of the BRAIN Initiative.

For instance, researchers funded by the BRAIN Initiative are developing imaging procedures to identify every connection in a mouse’s brain, or watch neurons fire in a circuit in real time (see page 315). Yet versions of these technologies that are suitable for use in humans are likely to be so expensive and difficult to use that they will be available only at a few research centres — and not at the average US hospital, much less in low-income countries.

Still, some researchers have found a sweet spot at which they can develop clinical imaging while working to understand the brain — even when working with limited equipment in

developing countries, Koroshetz says. Gretchen Birbeck, a neurologist at the University of Rochester in New York, is working in Zambia to understand how cerebral malaria can lead to epilepsy in children. She uses a magnetic resonance imaging (MRI) machine to watch how neural activity changes over time in the brains of children with malaria — and how, in some cases, epilepsy emerges.

Using existing technology, “we can really drill down and ask some important scientific questions that would have implications much more broadly”, Birbeck says. She hopes to do so despite using an MRI machine that is only about 10% as powerful as those in many US hospitals.

The WHO meeting participants also noted that existing technologies — such as mobile phones, which are prevalent in many developing countries where doctors are not — can be used to apply research findings to public health. Farrah Mateen, a neurologist at Massachusetts General Hospital in Boston who attended the WHO’s July meeting, has developed a cheap cap studded with electroencephalography sensors and an app that can identify the brain patterns that occur in people with epilepsy when they are between seizures. Her study of 205 people with epilepsy in Bhutan, now in review at a journal, showed that the app could reliably detect this neural activity — allowing health-care workers with little training to identify the type of epilepsy someone has and which drug could best treat it.

But one big question looms over the WHO’s push to ensure that brain-mapping projects yield clinical benefits: the agency does not fund research. Instead, it hopes to convert private donors and governments to its way of thinking. “We’re just going to try to influence the scientists as well as the funders to see whether there could be a better balance between long-term outcomes and more short-term, public-health-oriented outcomes,” says Saxena.

Not everyone is ready to reset their priorities. Tetsuo Yamamori, a neuroscientist at RIKEN Brain Institute in Saitama and vice project leader of the Brain/MINDS project, says that its primary goal will remain the same: linking behaviour to mapping brain activity in genetically engineered marmosets. But he agrees that developers of cutting-edge technologies should consider how they could be used in people. Yamamori adds that he will continue to participate in the WHO’s brain-mapping discussions.

Saxena is optimistic that the WHO’s lobbying will pay off. Eventually, he hopes to establish a network of researchers working to transform advances in brain research to improvements in medicine. “We really don’t understand this organ,” he says. “We need to make incremental advancements, and those need to be translated into something actionable in a short time.” ■

**“We really don’t understand this organ.”**



An ecosystem's ability to support crops is affected by the bacteria and fungi in the soil.

# ECOLOGY

# Mapping Africa's soil microbiome

*Sub-Saharan project could one day help ecosystems resist climate change and improve agriculture.*

BY SARAH WILD

One thousand ziplocked bags of soil from ten countries will form the basis of the first large-scale survey of the microbial life hidden underground in sub-Saharan Africa. The leaders of the African soil microbiology project hope that the data will one day help to drive better agricultural practices and to protect ecosystems and crops in the face of climate change.

"Soils are critical and soil health is vital for human and animal livelihoods," says Don Cowan, director of the Centre for Microbial Ecology and Genomics at the University of Pretoria in South Africa. He launched the project on 8 October at the consortium's first meeting in Pretoria.

Researchers increasingly recognize the importance of soil microbes to ecology and agriculture. Some bacteria and fungi colonize plant roots, promoting the plant's growth. A diverse population of soil microbes helps to regulate an ecosystem's climate, and maintains the fertility of the soil and its ability to support crops. And biotechnology companies

including Monsanto are testing additives that contain soil microbes for their ability to improve agricultural productivity.

Yet the diversity and sheer number of soil microbes — as well as the extent to which populations vary across the planet — is not well understood, says microbial ecologist Noah Fierer, who researches microbes in diverse environments at the University of Colorado at Boulder. A 2014 study of Central Park in New York revealed that the majority of its soil microbes were new to science. "We know surprisingly little about the soil microbiome," says Fierer.

This is particularly true for sub-Saharan Africa. South Africa, a scientific leader in the region, has a long history of soil science, but until now, its researchers have focused on soil chemistry rather than soil microbes. An understanding of the soil microbiome and its effect on crops is especially relevant now because microbial communities are expected

**"We know surprisingly little about the soil microbiome."**

to be affected by climate change.

Funded by the US Agency for International Development (USAID), the African project will take samples in South Africa, Namibia, Botswana, Zimbabwe, Mozambique, Zambia, Kenya, Ethiopia, Côte d'Ivoire, and Nigeria.

The goal of the three-year initiative is to create a broad survey of soil chemistry and microbiology across a range of regions and climates. Its researchers, led by Cowan, will use climate, topography and geology to choose 1,000 sample locations most likely to capture the maximum diversity. At each point, scientists will record data about the area, including temperature and altitude, and take photos of the site as well as digging the samples, which they will send back to Cowan's lab in South Africa. Each country is responsible for its own sample collection using a standardized method.

Cowan's team, in collaboration with researchers sent from partner countries, will extract the DNA from the samples — and amplify and sequence sections containing specific DNA tags that mark them out as bacteria. The researchers expect to find unknown bacterial genomes as well as known ones. In the project's next phase, they will look at soil fungi.

There are practical challenges ahead. Zimbabwe is under US sanctions and not eligible for USAID support, so Cowan must find alternative funding to undertake sampling there. And the project may run up against anti-biopiracy legislation, which restricts the movement of samples across borders. "If it proves to be impossible to get all the appropriate permits from partner nations, we'll just have to leave them out," says Cowan.

The project "represents a key step to charting the diversity of the soil microbiome and mapping the hidden biodiversity found below ground", says Fierer. Although valuable for creating a baseline for further studies, these kinds of data can provide only a superficial understanding of the microbiome, notes plant molecular biologist Simona Radutoiu at Aarhus University in Denmark. Her work involves in-depth studies of how soil bacteria interact with plants, which is necessary to understand the role of microbes in plant health. "This is a very good initiative," she says, "but is a start and should be continued, improved and sustained."

"The resolution is not great," acknowledges Cowan, "but the novelty is enormous." He adds: "We don't have the resources to do the coverage and analyses that would take the project to a more sophisticated level."

He hopes that the project will develop as more African scientists are trained to perform microbiome analyses, allowing more samples to be taken and a higher resolution of sampling. "We know our partners are struggling to build cutting-edge labs," he says. "We can't help them build their laboratories, but we can bring students here to train them in the practicalities." ■



## BIOMEDICAL RESEARCH

# Tracker flags failures to report trial results

*Computerized search of trial registry lists worst offenders.*

BY HEIDI LEDFORD

An automated tool has trawled through thousands of records on the world's leading clinical-trials database to reveal which drug firms and academic institutions are not publishing the results of their trials.

The failure to publish is already well documented: multiple studies have variously reported that 25–50% of clinical-trial results remain unpublished years after the trials are completed. But software such as the tool described in a paper published online in *F1000Research* (A. Powell-Smith and B. Goldacre, *F1000Research* <http://doi.org/bsnn>; 2016) on 3 November allows for a more comprehensive search than was previously possible, says corresponding author Ben Goldacre, a clinical-research fellow at the University of Oxford, UK.

(The publication has yet to be peer reviewed.)

Automating the process also means that results can be updated regularly, which keeps the pressure on trial sponsors who fail to report.

## COMPUTER CHECK

Goldacre and his Oxford-based co-author, web developer Anna Powell-Smith, developed the tool to search the ClinicalTrials.gov database for trials that were completed at least two years ago. The program attempted to match those trials with results published in that database or in the research repository PubMed.

Of nearly 26,000 trials evaluated, 45.2% had no published results. The team also built a website ([go.nature.com/2emchhz](http://go.nature.com/2emchhz)) that enables users to view clinical-trials sponsors in order of who is the best — or worst — at publishing their results. “If anyone wants to improve their

score or improve their ranking, all they have to do is publish their results,” says Goldacre.

Automated analyses are increasingly the norm for studies that scan for clinical-trial transparency, says Jennifer Miller, a medical ethicist at New York University's Langone Medical Center. She points to the Good Pharma Scorecard initiative launched by Bioethics International, a charity that Miller founded. The initiative ranks new drugs and companies on clinical-trial transparency on the basis of automatic analyses and machine learning. But, unlike with Goldacre's tool, the Scorecard team checks work manually and confirms findings with clinical-trial sponsors, she says.

Automating the search can lead to a sacrifice in precision, Goldacre acknowledges. For example, the search might miss published results if they are not tagged with a number assigned by the ClinicalTrials.gov database, or if the journal in which they are published is not listed in PubMed.

Despite some discrepancies in how individual studies were scored, Goldacre says, the trends from his tracker are similar to those previously published by manual surveys on smaller subsets of data. He hopes that the ability to regularly update results will incentivize trial sponsors to improve. “This is such a serious business,” he says. “We need to maintain the pressure.” ■



# HOW TO DEFEAT DEMENTIA



## The three things that could help prevent a meltdown in health-care systems worldwide.

BY ELIE DOLGIN

**T**here are not a lot of things that could bring together people as far apart on the US political spectrum as Republican Newt Gingrich and Democrat Bob Kerrey. But in 2007, after leading a three-year commission that looked into the costs of care for elderly people, the political rivals came to full agreement on a common enemy: dementia.

At the time, there were fewer than 30 million people worldwide diagnosed with the condition, but it was clear that the numbers were set to explode. By 2050, current predictions

suggest, it could reach more than 130 million, at which point the cost to US health care alone from diseases such as Alzheimer's will probably hit US\$1 trillion per year in today's dollars. "We looked at each other and said, 'You know, if we don't get a grip on Alzheimer's, we can't get anything done because it's going to drown the system,'" recalls Gingrich, the former speaker of the US House of Representatives.

He still feels that sense of urgency, and for good reason. Funding has not kept pace with the scale of the problem; targets for treatments



are thin on the ground and poorly understood; and more than 200 clinical trials for Alzheimer's therapies have been terminated because the treatments were ineffective. Of the few treatments available, none addresses the underlying disease process. "We're faced with a tsunami and we're trying to deal with it with a bucket," says Gingrich.

But this message has begun to reverberate around the world, which gives hope to the clinicians and scientists. Experts say that the coming wave can be calmed with the help of just three things: more money for research, better diagnostics and drugs, and a victory — however small — that would boost morale.

"What we really need is a success," says Ronald Petersen, a neurologist at Mayo Clinic in Rochester, Minnesota. After so many failures, one clinical win "would galvanize people's interest that this isn't a hopeless disorder".

### COST CALCULATIONS

Dementia is the fifth-biggest cause of death in high-income countries, but it is the most expensive disease to manage because patients require constant, costly care for years. And yet, research funding for dementia pales in comparison with that for many other diseases. At the US National Institutes of Health (NIH), for example, annual funding for dementia in 2015 was only around \$700 million, compared with some \$2 billion for cardiovascular disease and more than \$5 billion for cancer.

One problem is visibility. Other disease communities — most notably, people affected by breast cancer and HIV/AIDS — have successfully advocated for large pots of dedicated research funding. But "there simply wasn't any comparable upswell of attention to Alzheimer's", says George Vradenburg, chair and co-founder of UsAgainstAlzheimer's, a non-profit organization in Chevy Chase, Maryland.

The biggest reason, he says, is that "the victims of the disease hide out". Dementia mostly affects elderly people and is often misconstrued as a normal part of ageing; there is a stigma attached to the condition, and family care-givers are often overworked and exhausted. Few are motivated enough to speak up.

However, social and political awareness has increased in the past five years. "We all started to work together a lot more, and that helps," says Susan Peschin, chief executive at the Alliance for Aging Research in Washington DC, one of more than 50 non-profit groups in the Accelerate Cure/Treatments for Alzheimer's Disease coalition.

The impact can be seen in government investments. France took action first, creating a national plan for Alzheimer's in 2008 that included €200 million (US\$220 million) over five years for research. In 2009, the German Centre for Neurodegenerative Diseases in Bonn was created with a €66-million annual budget. And UK spending on dementia research more than doubled between 2010

and 2015, to £66 million (US\$82 million). The European Union has been dishing out tens of millions of euros each year for dementia studies through the Innovative Medicines Initiative and the Joint Programming process, and Australia is now about halfway through doling out its Aus\$200-million (US\$150-million), five-year dementia-research fund.

"This is a global challenge, and no one country will be able to solve the problem," says Philippe Amouyel, a neurologist and geneticist at the University Hospital of Lille in France. Yet it's the United States that has been the biggest backer by far, thanks in part to efforts by Gingrich and Kerrey. The NIH's annual budget for Alzheimer's and other dementias jumped in the past year to around \$1 billion, and there

## "WE'RE FACED WITH A TSUNAMI AND WE'RE TRYING TO DEAL WITH IT WITH A BUCKET."

is support for a target to double that figure in the next few years — even in the fractious US political landscape. "Alzheimer's doesn't care what political party you're in," says Kerrey.

Two billion dollars is "a reasonable number", says Petersen, who chairs the federal advisory board that came up with the target in 2012. Now, he adds, the research community just needs to work out "what are we going to do with it if in fact we get it?".

The answer could depend in large part on the fate of a drug called solanezumab, developed by Eli Lilly of Indianapolis, Indiana. This antibody-based treatment removes the protein amyloid- $\beta$ , which clumps together to form sticky plaques in the brains of people with Alzheimer's. By the end of this year, Lilly is expected to announce the results of a 2,100-person clinical trial testing whether the drug can slow cognitive decline in people with mild Alzheimer's. It showed preliminary signs of cognitive benefit in this patient population in earlier trials (R. S. Doody *et al.* *N. Engl. J. Med.* **370**, 311–321; 2014), but the benefits could disappear in this final stage of testing, as has happened for practically every other promising compound.

No one is expecting a cure. If solanezumab does delay brain degradation, at best it might help people to perform 30–40% better on cognitive tests than those on a placebo. But even such a marginal gain would be a triumph. It would show scientists and the drug industry that a disease-modifying therapy is at least possible. By contrast, another setback could bring recent momentum in therapeutic development to a halt.

"This is a fork in the road," says John Hardy, a neurogeneticist at University College

London. "This is going to be a very important outcome, way beyond the importance for Lilly and this particular drug."

On a scientific level, success for solanezumab could lend credence to the much-debated amyloid hypothesis, which posits that the build-up of amyloid- $\beta$  in the brain is one of the triggers of Alzheimer's disease. The previous failure of amyloid-clearing agents led many to conclude that plaques were a consequence of a process in the disease, rather than the cause of it. But those in favour of the amyloid hypothesis say that the failed drugs were given too late, or to people with no amyloid build-up — possibly those with a different form of dementia.

For its latest solanezumab trial, Lilly sought out participants with mild cognitive impairment, and used brain scans and spinal-fluid analyses to confirm the presence of amyloid- $\beta$  in their brains. Another company, Biogen in Cambridge, Massachusetts, took the same approach to screening participants in a trial of its amyloid-targeting drug aducanumab. Earlier this year, a 165-person study reported early signs that successfully clearing amyloid- $\beta$  with the Biogen therapy correlated with slower cognitive decline (J. Sevigny *et al.* *Nature* **537**, 50–56; 2016).

If those results hold up to further scrutiny, "that will at least tell us that amyloid is sufficiently upstream in the cascade that it deserves being targeted and tackled pharmacologically", says Giovanni Frisoni, a clinical neuroscientist at the University of Geneva in Switzerland who is involved in the drug's testing.

### TO DEFEAT, DELAY

Although debate over the amyloid hypothesis continues, interest is growing in earlier intervention with drugs that clear the protein. Reisa Sperling, a neurologist at Brigham and Women's Hospital in Boston, Massachusetts, worries that even mild dementia is a sign of irreparable brain-cell death. "You can suck all the amyloid out of the brain or stop it from further accumulating, but you're not going to grow those neurons back."

That is why she is leading Anti-Amyloid Treatment in Asymptomatic Alzheimer's, or A4, a \$140-million, placebo-controlled solanezumab study that aims to treat people with elevated amyloid levels before they show any signs of cognitive impairment. And A4 is not her only trial. In March, she and neurologist Paul Aisen of the University of Southern California's Alzheimer's Therapeutic Research Institute in San Diego launched a trial in 1,650 asymptomatic people with early signs of amyloid- $\beta$  build-up. It will test a pill from Johnson & Johnson that blocks  $\beta$ -secretase, an enzyme responsible for producing the toxic protein.

These interventions are known as secondary prevention because they target people who are already developing amyloid plaques. Sperling and Aisen also plan to test what's

called primary prevention. In August, they received NIH funding to start treating people who have normal brain levels of amyloid- $\beta$  and no signs of cognitive decline, but who have a high risk of developing Alzheimer's — because of a combination of factors such as age and genetics.

"The biggest impact we can have is in delaying the onset of the diseases," says David Holtzman, a neurologist at Washington University School of Medicine in St. Louis, Missouri, and an investigator in the Dominantly Inherited Alzheimer Network, which is testing the benefits of giving either solanezumab or another anti-amyloid therapy to people who inherit gene mutations that predispose them to develop Alzheimer's at an early age.

Secondary prevention could eventually mean screening everyone past middle age for signs of amyloid- $\beta$ , although the current testing methods are either expensive (\$3,000 brain scans) or invasive (spinal taps). Researchers have flagged a dozen possible blood-based biomarkers, but none has yet panned out, says Dennis Selkoe, a Brigham and Women's Hospital neurologist.

Yet a cheap and easy diagnostic test for amyloid- $\beta$  could ultimately prove unnecessary. In the same way that some have suggested giving cholesterol-lowering drugs to anyone at risk of heart disease, clinicians might eventually give anti-amyloid drugs to a broad set of people prone to Alzheimer's — even if they are not already amyloid positive, says Sperling.

## TARGET PRACTICE

Just as cholesterol is not the sole cause of heart disease, amyloid- $\beta$  is not the only driver of Alzheimer's. There's also tau, a protein that causes tangles in the brains of most people with Alzheimer's. Several pharmaceutical companies are targeting tau, but few large drug-makers have clinical candidates directed at other types of target. "They know how to modulate a specific target and keep looking under that lamp post, rather than venturing away from their comfort zones," says Bernard Munos, an industry consultant and former Eli Lilly executive.

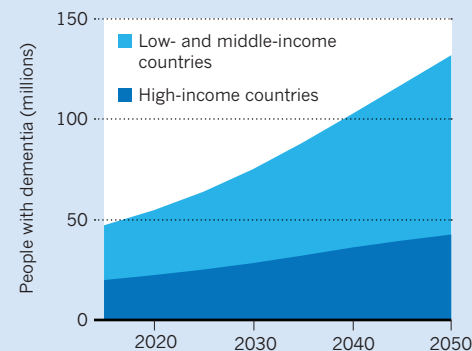
That's a problem, says Howard Fillit, chief science officer of the Alzheimer's Drug Discovery Foundation in New York City. "We really need to increase the diversity of targets we're tackling."

After amyloid and tau, the only target receiving much attention from researchers is neuroinflammation — the "third leg of the stool" in treating Alzheimer's, according to neurogeneticist Rudy Tanzi at Massachusetts General Hospital in Boston.

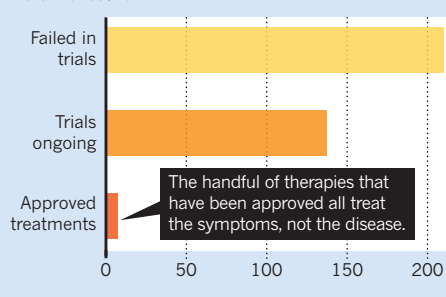
He likens Alzheimer's disease to a wildfire in the brain. Plaques and tangles provide the initial brush fires, but it's the accompanying neuroinflammation that fans the flames. Once the blaze is raging, Tanzi says, "putting out those

## THE APPROACHING WAVE

The number of people living with dementia worldwide will more than double in the next 35 years. Low- and middle-income countries will be the hardest hit.



Hundreds of clinical trials for Alzheimer's disease have been terminated because the treatments were ineffective.



brush fires that got you there isn't good enough."

This could explain why anti-amyloid drugs failed when given to people with full-blown dementia. For these individuals, perhaps reducing the inflammatory activity of brain immune cells called microglia could help. Drug researchers are now focusing on two genes, *CD33* and *TREM2*, that are involved in microglial function. But, says Tanzi, "there are two dozen other genes that deserve attention. Who knows if one of these new genes that no one is working on might lead to drug clues?"

## ALTERNATIVE AVENUES

Many Alzheimer's experts emphasize the need to develop better low-cost interventions that don't require drug research. At the University of New South Wales in Sydney, Australia, for example, geriatric psychiatrist Henry Brodaty is testing whether an Internet coaching tool that focuses on diet, exercise, cognitive training and mood can postpone disease development. "We know that two-thirds of the world's dementia is going to be in developing countries," he says (see "The approaching wave"). Lifestyle interventions, he argues, could be more broadly scalable than expensive drugs.

Researchers also need to look beyond Alzheimer's, to the many other types of dementia. Injuries to the vessels that supply blood to the brain cause a form called vascular dementia. Clumps of a protein called  $\alpha$ -synuclein underlie cognitive problems in

people with Parkinson's disease and also what's called Lewy body dementia. Tau deposits are often behind the nerve-cell loss responsible for frontotemporal dementia. And there are many other, equally devastating, drivers of serious mental decline.

"We should not be ignoring these other diseases," says Nick Fox, a neurologist at University College London, especially given that many types of dementia share biological mechanisms. Tackling one disease could help inform treatment strategies for another.

But perhaps the biggest hindrance to drug development today is more logistical than scientific, with clinical trials for dementia taking years to complete as investigators struggle to recruit sufficient numbers of study participants. "We need to get answers more quickly," says Marilyn Albert, director of the Johns Hopkins Alzheimer's Disease Research Center in Baltimore, Maryland.

One solution is trial-ready registries. By enrolling people who are interested in taking part in a study before it actually exists, investigators can start a trial as soon as a drug comes along for testing. "We have to register humanity in the task of defeating this disease," says Aisen.

The 1,600-person COMPASS-ND registry is being funded through the Canadian Consortium on Neurodegeneration in Aging. Member Serge Gauthier, a neurologist at McGill University in Montreal, says that finding participants can be challenging. But he adds that around one-third of the people who come to memory clinics such as his have what's known as subjective cognitive impairment — they might forget names or suffer from other "senior moments," but they do not meet the clinical definition of dementia.

They are perfect for trial-ready registries, says Gauthier: they are at an elevated risk of the disease, and they've demonstrated concern. Gauthier wants to find more people like them. He fits the profile himself, so he joined the Brain Health Registry, which has more than 40,000 participants so far and is led by researchers at the University of California, San Francisco. He takes regular cognitive tests, and could be asked to do more once potential diagnostic tools or therapies are ready for testing. "It's a fun thing to do," he says.

Voluntarily or not, people will need to face up to dementia, because in just a few short decades, pretty much everyone is going to have a friend or loved one affected by the disease. It's an alarming idea, and it should spur action, says Robert Egge, chief public policy officer of the Alzheimer's Association in Chicago, Illinois.

"We know where we're heading," he says. "The question is: are we going to get in front of it or not?" ■ [SEE BOOKS & ARTS P.166](#)

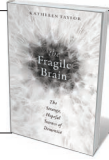
**Elie Dolgin** is a science writer in Somerville, Massachusetts.

SOURCE: TOP, ALZHEIMER'S DISEASE INTERNATIONAL; BOTTOM, B. MUNOS



# COMMENT

**CLIMATE** IPCC ocean report must focus on local variability **p.162**

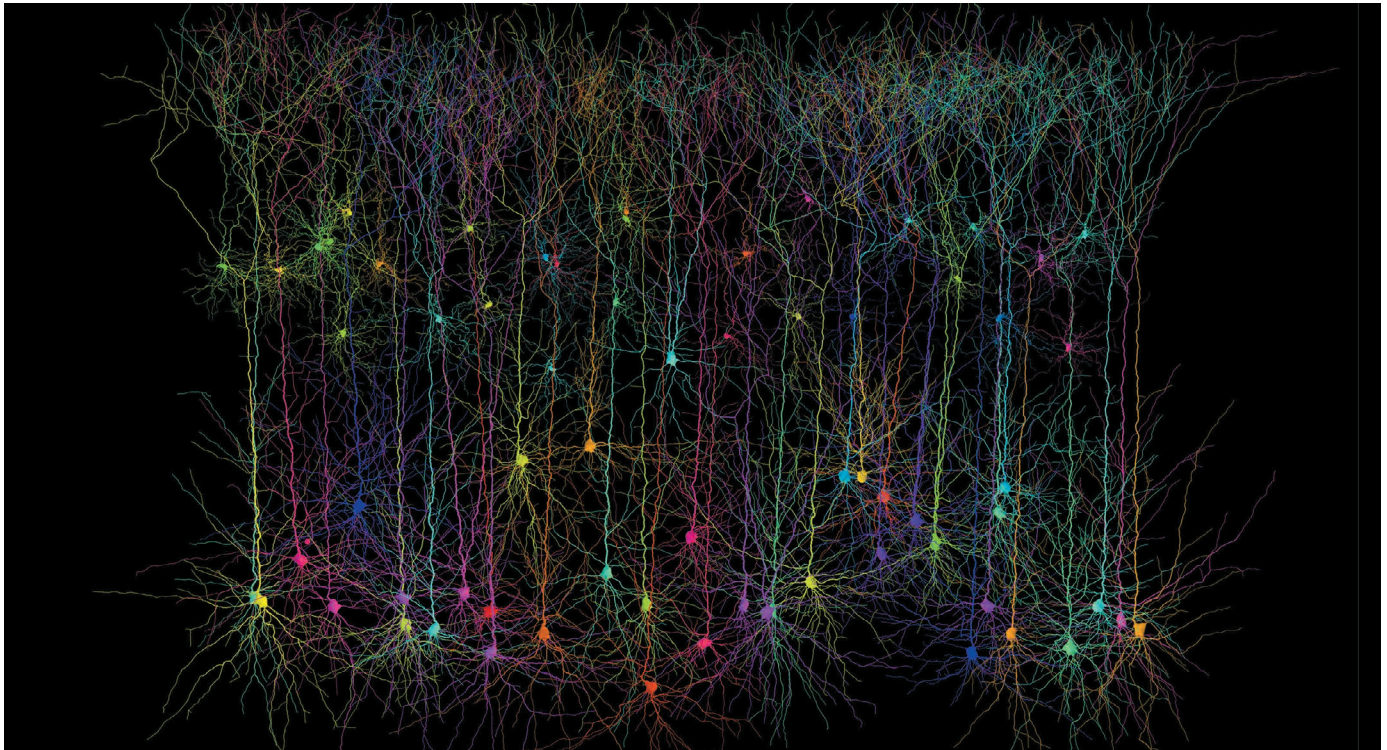


**NEUROSCIENCE** A sweeping survey of dementia, from research to care **p.166**

**METRICS** A call for China to step away from journal impact factor **p.168**

**PEER REVIEW** Should erstwhile co-authors always be ruled out? **p.168**

HERMANN CUNTZ; MICHAEL HÄUSSER



A network of pyramidal cells in the mouse cerebral cortex.

## A better way to crack the brain

Focused grass-roots collaborations that start small and scale up could overcome technical and sociological barriers to 'big' neuroscience, argue **Zachary F. Mainen, Michael Häusser and Alexandre Pouget.**

**A**t least half a dozen major initiatives to study the mammalian brain have sprung up across the world in the past five years. This wave of national and international projects has arisen in part from the realization that deciphering the principles of brain function will require collaboration on a grand scale.

Yet it is unclear whether any of these mega-projects, which include scientists from many subdisciplines, will be effective. Researchers with complementary skill sets

often team up on grant proposals. But once funds are awarded, the labs involved often return to work on their parts of the project in relative isolation.

We propose an alternative strategy: grass-roots collaborations involving researchers who may be distributed around the globe, but who are already working on the same problems. Such self-motivated groups could start small and expand gradually over time. But they would essentially be built from the ground up, with those

involved encouraged to follow their own shared interests rather than responding to the strictures of funding sources or external directives.

This may seem obvious, but such collaboration is stymied by technical and sociological barriers. And the conventional strategies — constructing collaborations top-down or using funding strings to incentivize them — do not overcome those barriers.

The European Commission's Human Brain Project (HBP), which ▶

► launched in 2013, involves more than 100 laboratories and the investment of at least US\$49 million per year. Following a reorganization, the HBP now emphasizes tool development, leaving the collective effort of gathering and analysing the data to other initiatives planned or under way, such as Japan's Brain/MINDS project, the US BRAIN initiative or China's proposed programme. It has been questioned how any of these will result in more than the sum of their parts<sup>1,2</sup>.

Of all the current big-neuroscience initiatives, perhaps the most effective has been that of the Allen Institute for Brain Science, a private non-profit organization in Seattle, Washington. More than 100 scientists working together at the institute have produced useful resources, including brain-wide maps of gene expression in mouse and human and, most recently, maps of neural activity in the mouse visual cortex<sup>3</sup>. But it is not clear how the Allen Institute's model — industrial processes in a centralized corporate organizational structure — could be applied to geographically distributed collaborations targeting more complex problems.

Some sceptics point to the teething problems of existing brain initiatives as evidence that neuroscience lacks well-defined objectives<sup>1,2,4</sup>, unlike high-energy physics, mathematics, astronomy or genetics.

In our view, brain science, especially systems neuroscience (which tries to link the activity of sets of neurons to behaviour) does not want for bold, concrete goals. Yet large-scale initiatives have tended to set objectives that are too vague and not realistic, even on a ten-year timescale.

## THE CHALLENGE

Several advances over the past decade have made it vastly more tractable to solve fundamental problems such as how we recognize objects or make decisions.

Researchers can now monitor and manipulate patterns of activity in large neuronal ensembles, thanks to new technologies in molecular engineering, microelectronics and computing. For example, a combination of advanced optical imaging and optogenetics can now read and write patterns of activity into populations of neurons<sup>5</sup>. It is also possible to relate firing patterns to the biology of the neurons being recorded, including their genetics and connectivity.

Other tools coming online include powerful statistical techniques for analysing data and advances in machine learning. There is also now a rich set of theories stemming from progress in fields such as statistical physics and computer science. Computational approaches have been used to understand, for instance, how neurons



The industrial approaches used to produce neuroscience resources at the institute founded by philanthropist Paul Allen may not work elsewhere.

in the retina and visual cortex encode information about visual scenes<sup>6,7</sup>.

But the experiments now possible are increasingly resource-intensive. The neuronal activity driving a simple behaviour, such as a mouse navigating a maze, could involve the cooperation of several hundred brain areas. Mapping the whole picture implies making recordings in many neurons from each area. Yet a typical 1–3-year study involves recording from relatively small populations of neurons in just a single area of the brain. And, as we will discuss, these data cannot at present be combined across labs.

Most new approaches for the collection and analysis of neural data require training and expertise across a range of domains — from genetics to optics to computational neuroscience. As in most disciplines, neuroscientists in one laboratory — let alone one scientist — rarely hold the entire set of requisite skills. Moreover, because labs do not normally share raw data, the fruits of difficult experiments cannot be fully exploited by groups with complementary expertise.

In short, a generation ago, neuroscientists were largely limited by theory and tools. Today, the bigger problem is effectively harnessing, as a community, what is already available.

## A SOLUTION

We propose that researchers join forces in 'meso-scale' collaborations of around 20 principal investigators and between 50 and 100 researchers to conduct experiments that are beyond the reach of single labs. Even at this scale, there will be many hurdles to clear. Specifically, an effective collaboration would need to do the following.

**Focus on a single brain function.** The downfall of many neuroscience collaborations — and especially of mega-projects — is setting goals that are too broad. The common goal has to be ambitious, yet reachable within, say, ten years, and well defined. A whole-brain theory of one brain function — a single behaviour — could meet those requirements. If a collaboration were largely limited to labs interested in the same behaviour — such as courtship in fruit flies, or foraging in mice — clear, shared objectives could be defined at the start. The labs would apply a range of recording and manipulation techniques to the same common behavioural task, allowing the functional data to be seamlessly combined.

To assemble a team of experts on such a focused problem, a collaboration would need to incorporate participants distributed throughout the world. In the past, physical proximity was indispensable for effective interaction. Now, online collaboration tools — such as Slack, GitHub or Google Docs — have changed the game. Scientists must harness these to plan experiments, make decisions, discuss problems and more. For more specific needs, new tools may need to be invented.

## Combine experimentalists and theorists

There is a growing consensus that theory is indispensable for grappling with brain complexity. In particular, theory is essential for making and testing predictions about how observations made at the cellular or circuit level will relate to those made at the behavioural level.

Yet, the challenges of implementing modern experimental approaches and of getting to grips with the mathematical language of theory mean that neuroscientists



still tend to become either experimentalists or theorists. Moreover, whole labs are typically either experimental or theoretical. Theorists and experimentalists often meet at conferences to share ideas. They rarely converge when it comes to the design and interpretation of experiments.

So, concrete steps are required to catalyse more meaningful interactions, such as embedding theory PhD students and post-docs in experimental labs and vice versa.

**Standardize tools and methods.** Neuroscientists frequently live on the ‘bleeding’ edge technologically, building bespoke and customized tools. This do-it-yourself approach has allowed innovators to get ahead of the competition, but hampered the standardization of methods essential to making experiments efficient and replicable.

Remarkably, it is standard practice for each lab to custom engineer all manner of apparatus, from microscopes and electrodes to the computer programmes for analysing data. Thousands of labs worldwide use the calcium sensor GCaMP, for example, for imaging neural activity *in vivo*. Yet neither the microscopes used for GCaMP imaging nor the algorithms used to analyse the resulting data sets have been standardized.

The data sets generated by a functional neuroscience experiment are large. They can also be complex and multimodal in ways that, say, genomic data might not be, embracing recordings of activity, behavioural patterns, responses to perturbations, and subsequent anatomical analysis. Researchers have no agreed formats for integrating different types of information. Nor are there standard systems for curating, uploading and hosting highly multimodal data. Recently, initiatives such as Neurodata Without Borders ([www.nwb.org](http://www.nwb.org)) have finally started to address this.

Worse, neuroscientists lack standardized vocabularies for describing the experimental conditions that affect brain and behavioural functions. Such a vocabulary is needed to properly annotate functional neural data. For instance, even small differences in when a water drop is released can affect how a mouse’s brain processes this event, but there is no standard way to specify such aspects of an experiment.

**Share data.** To maximize the effectiveness of a collaboration, all the data collected would need to be communal across the entire group. This may seem obvious, but most of the data currently being collected using high-throughput imaging and recording techniques are still effectively inaccessible to anyone outside the labs doing the studies. Journal requirements to make data public and efforts to build public databases and standardize

formats have had little effect across much of neuroscience.

Beyond standardization problems, there are also substantial disincentives to data sharing that must be addressed by grassroots collaborations. Sharing can yield big common benefits, allowing data sets from multiple laboratories to be combined and theorists to test their ideas. But a lab risks losing out to competitors if its generosity is unreciprocated.

An effective collaboration must create the technical means to share, and engender a sphere of trust within which it is safe to do so. The principle of sharing — of data, resources and plans — would need to be agreed as a precondition of joining a collaboration, and effectively enforced.

**Assign credit in new ways.** Like many other areas of biomedicine, neuroscience is dominated by a competitive and individualistic culture. Indeed it is largely this culture that hinders standardization and cooperation. The Human Genome Project opened up a more cooperative attitude towards data in the field of genetics<sup>8</sup> that has been reverberating ever since<sup>9</sup>. But the intricacy of what we are proposing — the complex coordination of experiments and immediate sharing of raw data — goes well beyond most open-science norms and will be challenging.

To jump-start a culture of collaboration along the lines we are envisaging, groups of established scientists, who have less career pressure, could lead the way. The

**“Neuroscience is dominated by a competitive and individualistic culture.”**

graduate students and postdocs involved in collaborations might need other ways to earn recognition than the current

standard — being the first author on a paper. Neuroscience can take inspiration from fields such as particle physics, where these issues have been faced for years. And initiatives such as the Contributor Roles Taxonomy are revamping how contributions to research are defined and recognized ([casrai.org/credit](http://casrai.org/credit)).

## SMALL STEPS

Elements of our proposal have been discussed before, and the problems it addresses are not new. So why is now the time to finally make it work? First, the forces that drive neuroscientists apart, especially competition for resources, are stronger than ever. Second, advances in technology for information sharing, such as cloud services, are only now making distributed collaboration feasible. Third, because of the wealth of new experimental and theoretical tools, the potential benefits

of collaboration to many may at last outweigh the risks to individuals.

If grounded in the same principles that make small-scale collaboration so successful — including equality and transparency — medium-sized collaborations would be fundamentally different from the much looser networks and top-down initiatives typically associated with big science. Doubts might be raised about how far such groups could be scaled up. But just such a distributed model, avoiding a central command-and-control structure, characterizes one of the largest and most effective mega-projects<sup>10</sup> — the ATLAS collaboration at CERN, Europe’s particle-physics lab near Geneva, Switzerland.

Effective large-scale neuroscience initiatives may be the goal, but they cannot be created from scratch even with vast funds. Yet there are many useful things that funders can do to help. First, support multiple medium-sized collaborations that set concrete goals (rather than crowning single mega-projects) and keep renewing those that demonstrate progress. Second, encourage investigators to combine individual grants to join collaborations. Third, underwrite the less sexy but crucial aspects of collaboration, such as management and support personnel, that are otherwise difficult to fund. Fourth, fund the development of collaborative-science software, which would offer great returns on investment. Fifth, study, experiment with and support new ways of assigning credit that promote cooperation.

Small but carefully considered steps, not grand gestures, will ensure that much-needed neuroscience collaborations take root and flourish. ■

**Zachary F. Mainen** is director of research at the Champalimaud Centre for the Unknown, Lisbon, Portugal. **Michael Häusser** is professor of neuroscience at University College London, UK. **Alexandre Pouget** is professor of neuroscience at the University of Geneva, Switzerland. e-mail: [zmainen@neuro.fchampalimaud.org](mailto:zmainen@neuro.fchampalimaud.org)

1. Hunter, P. *EMBO Rep.* **16**, 685–688 (2015).
2. Poo, M.-M. *Natl Sci. Rev.* **1**, 12–14 (2014).
3. Hawrylycz, M. et al. *Proc. Natl Acad. Sci. USA* **113**, 7337–7344 (2016).
4. Frégnac, Y. & Laurent, G. *Nature* **513**, 27–29 (2014).
5. Emiliani, V., Cohen, A. E., Deisseroth, K. & Häusser, M. *J. Neurosci.* **35**, 13917–13926 (2015).
6. Tkačik, G. et al. *PLoS Comput. Biol.* **10**, e1003408 (2014).
7. Moreno-Bote, R. et al. *Nature Neurosci.* **17**, 1410–1417 (2014).
8. Green, E. D., Watson, J. D. & Collins, F. S. *Nature* **526**, 29–31 (2015).
9. Yozwiak, N. L., Schaffner, S. F. & Sabeti, P. C. *Nature* **518**, 477–479 (2015).
10. Boissot, M. *Collisions and Collaboration: The Organization of Learning in the ATLAS Experiment at the LHC* (Oxford Univ. Press, 2011).



Whitehaven Bay in Australia's Great Barrier Reef.

# Forecast ocean variability

The IPCC should supply policymakers with realistic regional projections of how the seas will respond to warming, write **Daniela Schmidt** and **Philip W. Boyd**.

**T**he ocean modulates Earth's climate and provides us with food, coastal protection, clean seawater and oxygen.

Only the latest assessment of the Intergovernmental Panel on Climate Change (IPCC)<sup>1,2</sup>, in 2014, had dedicated chapters on the oceans. Now the IPCC is preparing a special interdisciplinary report on the ocean and cryosphere (Earth's snowy and icy regions). Next month, a group of scientists will decide what to include in the report, which will be published in 2019.

Offering robust projections that can be translated into practical policy is central. The report must join up with the United Nations' Sustainable Development Goals. For example, goal 14 tasks governments to "sustainably manage and protect marine

and coastal ecosystems from land-based pollution, as well as address the impacts of ocean acidification". The report must help marine managers to make decisions here and now.

The IPCC needs to shift its approach. It must offer both short-term climate-change projections and longer-term ones, and acknowledge the variable nature of the oceans — not just global average trends. Its report must include: forecasts of how fluctuations and shifts in surface temperatures and pH are driven by both natural and anthropogenic climate change; near-term predictions of extreme conditions such as marine heat waves on regional scales; and the biological mechanisms that underpin how key organisms, and hence important ecological systems, respond to climate change.

This will take the IPCC out of its comfort zone. Decadal projections and regional foci represent greater uncertainties. But such information is necessary to safeguard our seas.

## NOISY WATERS

The impacts of climate change on the oceans are usually depicted using graphs. Lines represent projections of long-term globally averaged quantities such as relentless rises in mean sea surface temperature or acidification. But the real ocean is noisy. Its conditions simultaneously undergo fast and slow variations as well as local, regional and global ones.

It is important to quantify the long-term average state of the ocean. Eventually, the influence of anthropogenic climate change will be larger than that of ongoing natural variability<sup>3</sup>. This transition is known as the emergence. But we are not there yet. The present oceanic signature of anthropogenic climate change is still comparable to, and thus difficult to disentangle from, natural and regional climate variability such as the El Niño Southern Oscillation, cycles in winds and sea surface temperatures over the tropical east Pacific Ocean.

Emergence will happen at different times in different places. For example, the tropics are already recording extreme temperatures, whereas the emergence is several decades away at mid-latitudes<sup>4</sup>.

Natural climate variability can offset or amplify climate change trends temporarily (see 'Reading the waves'). For example, an apparent<sup>5</sup> slowing or 'hiatus' in global average temperature rise between 1998 and 2012 led some critics to downplay anthropogenic climate change. Natural variability also reflects more extreme conditions, such as latest strong El Niño warming event.

As anthropogenic climate change increases, periods of extreme conditions<sup>6</sup> are expected to become more frequent, severe and lengthy. These will have adverse effects on marine ecosystems<sup>7</sup>. For example, in 2011 the west coast of Australia encountered sea surface temperatures that were 2–4 °C warmer than average for 10 weeks. Its kelp forest, usually 800 kilometres long, shrank by 43%<sup>7</sup>.

These fluctuations are confusing for marine-resource managers, policymakers and the public. They make decisions about how best to adapt to climate change difficult, and short term forecasts unreliable.

## LOCAL ACTIONS

Oceanic and atmospheric processes are linked. So long-distance connections between regional climate patterns also confuse local marine measures and predictions.

Policymakers and marine managers need to know more about this variability and its impacts. Regional and local scales are most



SOURCE: A, REF. 3; B, NOAA, EQUATORIAL PACIFIC SEA SURFACE TEMPERATURES, [HTTP://GO.NATURE.COM/2F9GWG9](http://go.nature.com/2F9GWG9)

pertinent to managing marine resources. There will be hotspots of change, such as sites of marine heat waves<sup>7</sup>, or places where regional warming exceeds the global average, such as the western Antarctic Peninsula. Yet limits to the resolution and boundary conditions of global circulation models make it hard to represent changes in coastal regions. Regional projections from global climate models rarely agree and they exclude other human stressors such as fishing pressure and pollution.

The IPCC special report needs to tease apart how combinations of global, regional and local stressors will increase pressures on marine ecosystems and services in particular places. This would help local managers to buy time to mitigate the combined effects of multiple factors. For example, managing the run-off of sediment, nutrients and contaminants into coastal waters near Australia's Great Barrier Reef (see, for example, [go.nature.com/2ex5leq](http://go.nature.com/2ex5leq)) should give corals respite from devastating outbreaks of crown-of-thorns starfish (*Acanthaster planci*) that can add to the regular stress of bleaching (expelling algae under warmer conditions) during El Niño events.

Developing regional and local marine policies requires better understanding of governance mechanisms, management and trade practices too. The IPCC report should include examples of using local know-how to underpin policies from across the three IPCC working groups (such as the Great Barrier Reef example). The focus is still too much on physical signals of climate change<sup>8</sup>. Interdisciplinary studies need to be done on the legal and economic frameworks that support regional resilience — social and ecological.

## LIFE STORY

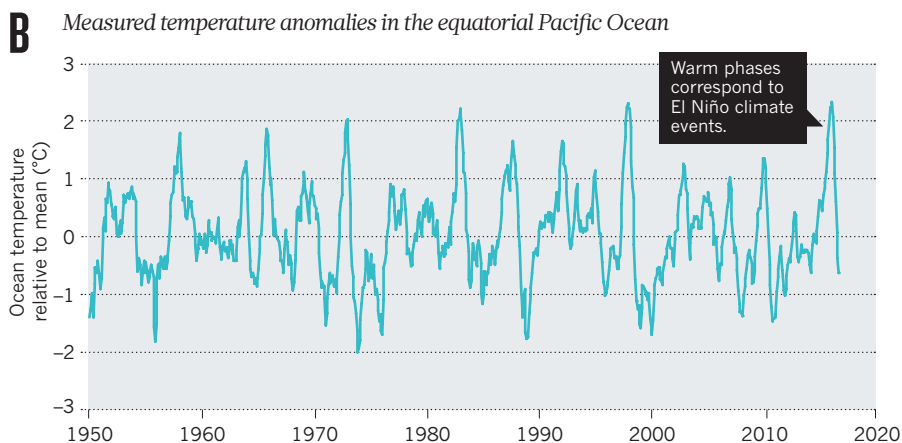
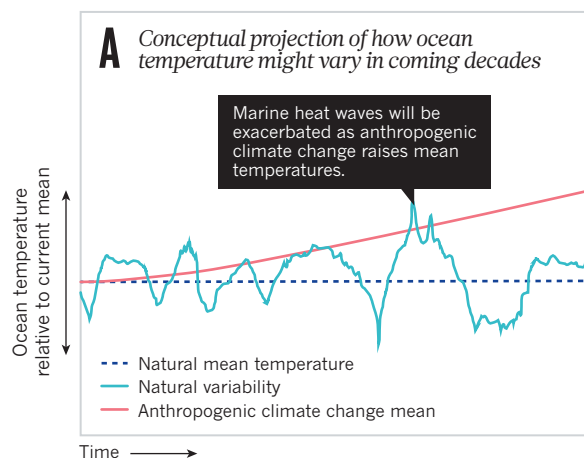
A final challenge for ocean scientists is to describe how marine life in diverse ecosystems will respond to the complex matrix of anthropogenic change<sup>8</sup>. Biologists do not understand fully the cumulative responses of the key components of ecosystems to a changing climate. What we have now are snapshots of how a few species within coastal food webs react to more acidic or warmer conditions.

Organisms may react in a nonlinear way<sup>1</sup>. If a species is already living in the warmest conditions it can handle, any further temperature rise will have lethal consequences, whereas a cooling would improve fitness. For example, many species of plankton in the tropical ocean are thought to be close their upper temperature limits<sup>8</sup>.

Experiments need to reflect the wider range of changes to local ocean conditions that will occur in the coming years and decades. And researchers need to consider how extremes and fluctuating conditions

## READING THE WAVES

Water temperatures fluctuate naturally around mean conditions as well as regionally. Year-to-year or decadal variations are currently hard to distinguish from global warming trends (A). Different places will encounter varying marine and climate pressures at different times. Data from the Pacific Ocean, for example, reveal many rapid changes in local sea temperatures (B).



affect physiologies. For example El Niño events may increase the mortalities of some species in the Pacific by adding to anthropogenic warming; whereas the cooler La Niña phase of the cycle would offer respite. How these processes balance out, if the periods of relief are long enough to allow recovery, and which species will be most affected are all open questions. This uncertainty is fundamental to our ability to predict the societal impacts of these ecological changes.

Places where warming is now above the global average are natural laboratories<sup>9</sup>. They include marine sanctuaries such as the Galapagos Islands and areas where humans rely heavily on ocean resources, such as southeast Asia and western Africa<sup>3</sup>. Environmental impact assessments in these places have revealed that some ecological changes, such as the loss of Australian kelp, are irreversible even if the physical environment returns to average conditions<sup>7</sup>. During heat waves, for instance, warmer water species migrate into colder water habitats, where they may replace the endemic species. Extreme events therefore can push ecosystems past tipping points.

Describing the oceans' variability will ensure that the IPCC report builds a bridge to the sustainable development goals. This

must be reflected in the choices made in December's meeting of ocean experts. ■

**Daniela Schmidt** is professor of palaeobiology, School of Earth Sciences, University of Bristol, UK. **Philip W. Boyd** is a marine biogeochemist, Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Australia.  
 e-mail: [d.schmidt@bristol.ac.uk](mailto:d.schmidt@bristol.ac.uk)

1. Pörtner, H. O. et al. in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds C. B. Field et al.) Ch. 6, 411–484 (Cambridge Univ. Press, 2014).
2. Hoegh-Guldberg, O. et al. in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds V. R. Barros et al.) Ch. 30, 1655–1731 (Cambridge Univ. Press, 2014).
3. Mora, C. et al. *Nature* **502**, 183–187 (2013).
4. Hawkins, E. et al. *Nature* **511**, E3–E5 (2014).
5. Xie, S.-P. *Nature Clim. Change* **6**, 345–346 (2016).
6. IPCC. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change* 582 (Cambridge Univ. Press, 2012).
7. Wernberg, T. et al. *Science* **353**, 169–172 (2016).
8. Boyd, P. W. et al. *Glob. Change Biol.* **22**, 2633–2650 (2016).
9. Hobday, A. J. & Pecl, G. T. *Rev. Fish Biol. Fisher.* **24**, 415–425 (2014).



Vincente, 87, has dementia; he is cared for at home by his daughter.

## NEUROSCIENCE

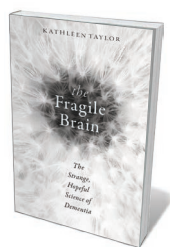
# Tide of forgetting

**Michael Heneka** applauds a sweeping survey of dementia that explores research, diagnosis and care.

**T**he global rise in dementia should surprise no one. The figures — such as the 9.9 million new diagnoses each year — have been known for decades. As slow as we are to accept such vast changes on a personal, societal and political level, so research is slow to uncover why our brains become fragile with age.

Neuroscientist and writer Kathleen Taylor's *The Fragile Brain* is about that research. But it is much more than a simple reflection on the best published hypotheses. Taylor has crafted a personal, astonishingly coherent review of our current state of knowledge about the causes of Alzheimer's disease and dementia, as well as possible solutions, from lifestyle adjustments to drug developments.

Filled with elegant metaphors, her study covers the detail of molecular biology and larger-scale analysis, including epidemiological observations and clinical studies. It extends to dementia due to multiple sclerosis, stroke and encephalitis. For instance, some 5–30% of people who have a first stroke develop dementia. But the book's focus is Alzheimer's disease, and rightly so: it is what up to 80% of people with dementia



**The Fragile Brain:**  
**The Strange, Hopeful Science of Dementia**

KATHLEEN TAYLOR  
Oxford University Press: 2016.

are diagnosed with. Taylor begins with a shocking juxtaposition, setting the costs of age-related disorders and of dementia alongside the scarcity in funding. In Britain, Australia and the United States, for example, funding for dementia research is a fraction of that for cancer — in the United States, just 18%. She contextualizes with reflections on the history of dementia research, deftly unravelling the roles of pioneering scientists Alois Alzheimer, Franz Nissl and Emil Kraepelin in describing the condition.

She then walks the reader through different brain elements and regions, from single neurons, transmitters and receptors to their complex interactions and function in cognition and behaviour. She describes the functional and structural loss of synapses as well

as the many ways in which amyloid folds and abnormally accumulates in our brains.

Taylor's discussion of risk is especially rich. She explains how to interpret studies on risk factors for age-related dementia and Alzheimer's, such as carrying the *APOE4* gene mutation. She emphasizes the influence of immune cells and their actions, such as clearing amyloid from the brain. She succinctly buries the long-fostered concept of the brain's immune privilege. And she isolates prime suspects in mediating further risks, from inflammation to the consequences of infection, vascular changes and production of reactive oxygen species such as hydrogen peroxide.

The big modifiable factors in early cognitive decline are, we now know, diabetes, obesity and smoking. Taylor's discussion here is sympathetic rather than hectoring, acknowledging the difficulties in shifting lifestyles. Likewise, her evidence on the influence of diet, food quality and habits is carefully handled, distilled to writer Michael Pollan's dictum: "Eat food. Not too much. Mostly plants."

Factors harder to influence by ourselves — such as stress, depression, environmental pollution or poverty — she frames as needing consideration in the larger social context. She discusses interventions to delay or prevent dementia, such as physical fitness and good education. We know from studies that these

can buffer genetic disadvantages such as carrying *APOE4*.

Taylor does not glide over the limitations of current interventions,

**NATURE.COM**

For more on science  
in culture see:

[nature.com/  
booksandarts](http://nature.com/booksandarts)

ANZENBERGER/EYEVINE



such as drugs that at best offer transient improvement. Despite decades of intensive research on dementia, a therapeutic breakthrough is yet to come. Taylor counters gloom by revealing avenues of current and future research, for instance further analysis of molecular pathways.

Taylor's discussion of hypotheses about the nature of dementia points to the limitations of concepts such as the physiological function of the amino acids in amyloid plaques in the brain. She offers a forthright assessment of the amyloid-cascade hypothesis and its pros and cons — for example, evidence of genetic mutations and lack of therapeutic success, respectively — and examines it in the context of ideas about the role of inflammation in the brain, particularly in Alzheimer's.

She is critical of how dementia research is conducted, discussing, for instance, the limitations of mouse models in analysing disease mechanisms. Her constructive suggestions include not restricting animal models to studies of the brain, but rather looking at the animal as a whole, because peripheral events such as infection have a neurological impact. However, as she notes, such moves would demand significant funding increases.

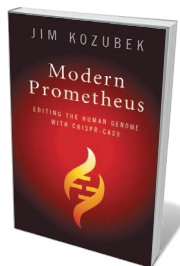
Finally, *The Fragile Brain* addresses how people interact with those who have dementia, and how both information and respect are needed. This is of utmost importance: information can quell many fears.

Taylor has the talent to make complex biology easy, but not trivial. There are many gems, such as her comparison of the immune system's roles of "peacekeeping" and healing with the US Marshall Plan to aid Western Europe after the Second World War. She nicely avoids over-interpretation of findings where research has not yet progressed beyond guesswork. Her provocative questions about genetics, habits, attitudes and levels of knowledge should prompt readers to reason and hypothesize for themselves, and to learn — for instance, whether a particular type of surgery and a particular anaesthetic act together to induce or aggravate dementia.

*The Fragile Brain*, as a trove of accessible, up-to-date science, has something to offer caregivers and families of people with dementia; medical professionals in dementia diagnosis and treatment; and specialists in memory clinics. It calls for change that reaches to the foundations of our society, suggesting that we adapt lifestyles, workplaces and institutions "to prioritize not labour and cash, but health and well-being". Some of this change has begun. We had better keep going. ■ **SEE NEWS FEATURE P.156**

**Michael Heneka** is a clinical neuroscientist in the Department of Neurodegenerative Disease and Gerontopsychiatry at the University of Bonn, Germany.  
e-mail: michael.heneka@ukb.uni-bonn.de

## Books in brief



### Modern Prometheus: Editing the Human Genome with Crispr-Cas9

Jim Kozubek CAMBRIDGE UNIVERSITY PRESS (2016)

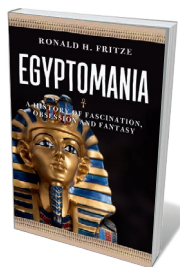
As the CRISPR-Cas9 method for rewriting genomes revolutionizes biotechnology, bioinformatician Jim Kozubek steps into the maelstrom with a weighty exploration of its discovery and implications. It should be noted that Kozubek was recently affiliated with an institution battling for patent rights to CRISPR gene-editing. His tome is also in need of a heavy edit. But he usefully pushes the discussion beyond obvious designer-baby concerns to the technique's limitations, and its broader implications for agriculture and the commercialization of science.



### Vertical: The City from Satellites to Bunkers

Stephen Graham VERSO (2016)

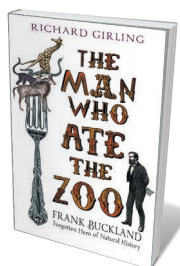
Our view of cities is perilously partial, argues urban geographer Stephen Graham. Mustering evidence in engineering, sociology and beyond, he argues for a new, vertical perspective — satellite to sewer — to reflect today's "intensified urban stacking". Seeing cities as Gordian knots of geopolitics, he gathers an impressive range of case studies to bolster his analysis. These compel and convince, from Saudi Arabia's high-rise vanity projects to Rio de Janeiro's *favelas* — which struggle with basic services beneath cable cars full of tourists — and the ultradeep mineral mines that service urban infrastructures.



### Egyptomania: A History of Fascination, Obsession and Fantasy

Ronald H. Fritze REAKTION (2016)

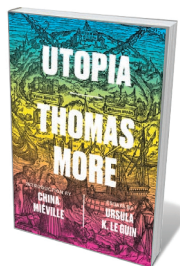
The richness, distinction and diversity of ancient Egyptian culture has fired imaginations for millennia. Here, historian Ronald Fritze examines 'Egyptomania' in detail and through time. As Herodotus and other classical scholars extolled Giza's pyramids and the great lighthouse at Alexandria, Egyptian cults and esoteric tracts seeped into Greece and Rome — to later fascinate and befuddle medieval and Renaissance scholars. The cracking of hieroglyphs, discovery of Tutankhamun's tomb and "mummymania" from the nineteenth century onwards ensured that the craze persists almost unabated today.



### The Man Who Ate the Zoo

Richard Girling CHATTO & WINDUS (2016)

Victorian zoologist and surgeon Frank Buckland occupies a peculiar place in science history. Like his renowned naturalist father William, he was both a serious researcher — rubbing shoulders with scientific heavyweights such as Michael Faraday — and an eccentric who dined on giraffe and panther. In this lively biography, Richard Girling revels in Buckland's phenomenal drive to master animal biology in a number of contexts: domestic menageries featuring marmots, a meerkat and a bear; a flood of natural-history writing; stints as a zoo medic; and distinguished contributions to fisheries science.



### Utopia

Thomas More VERSO (2016)

Five hundred years ago, English humanist Thomas More — who counselled Henry VIII and was executed at his order — published *Utopia*, a radical imagining of a society free from tyranny and suffering. This special edition is bookended by pieces from science-fiction greats Ursula Le Guin and China Miéville. In one essay, Le Guin suggests we edge back into history to see forward to a liveable future. How better than to revisit More, whose view of the social order as a "conspiracy of the rich" and war as inhuman resonate so powerfully today? **Barbara Kiser**

# Correspondence

## Revise rules on conflicts of interest

We contend that definitions of conflicts of interest (COI) in peer review need to be reassessed to reflect modern research practices. This could markedly increase the speed and quality of peer review.

For example, many potential reviewers are disqualified under current rules on co-authorship. However, research papers now have increasing numbers of co-authors and their interaction may be little more than episodic, with no genuine COI in practice. The judgement of an author who, say, contributed a data set to a paper is unlikely to be corrupted when reviewing a new paper from former co-authors.

In our editorial experience, co-authors typically have a sound understanding of each other's work and provide frank and constructive feedback. Using them as reviewers avoids settling for candidates who may be too far removed from the topic or not sufficiently senior in the field.

We suggest that only long-running co-authorship should be counted as a COI in peer review. Other potential COIs should include supervisor–student relationships, shared institutional affiliations and collaborators working on the same project, with an expiry date if appropriate.

**Indrė Žliobaitė, Mikael Fortelius** *University of Helsinki, Finland.*  
[indre.zliobaite@helsinki.fi](mailto:indre.zliobaite@helsinki.fi)

## Preserve specimens for reproducibility

The description of a new species without a preserved type specimen has always been permitted (T. Pape *et al.* *Nature* **537**, 307; 2016) — but it should not become the norm. Original specimens allow testing of the hypotheses that underlie descriptions and so ensure reproducibility — an obligation

and cornerstone of the scientific method.

It is taxonomic convention when describing a new species to deposit type specimens in a publicly accessible collection. This allows independent re-examination, reinterpretation and re-evaluation (*Nature* **535**, 323–324; 2016). Although photographs can point to possible undescribed species and help to document biodiversity, they are open to misinterpretation (and also to manipulation).

Photographs alone should remain the exception, used only when specimens cannot be preserved for technical, legal or conservation reasons. Properly vouchered specimens are otherwise essential in biodiversity research, just as “laboratory notebooks and records must be available for independent review” in the experimental sciences (C. G. Begley *et al.* *Nature* **525**, 25–27; 2015).

**Frank T. Krell\*** *Denver Museum of Nature & Science, Colorado, USA.*  
[frank.krell@dmns.org](mailto:frank.krell@dmns.org)  
\*On behalf of 5 correspondents (see [go.nature.com/2fjehxz](http://go.nature.com/2fjehxz) for a full list).

## Open up research evaluation in China

Strong academic opposition has led China's ministry of education to suspend a policy that would effectively control where Chinese researchers should publish their work. In my view, this scientifically disruptive intervention should never be reactivated.

The policy, launched by the ministry's Academic Degrees and Graduate Education Development Center (CDGDC) in April this year and repealed two weeks later, centred on an ‘A-list’ of top international journals. The position of a journal in the list is determined by impact factor, years after the global movement away

from its well-documented deficiencies as a tool for research assessment (see [www.ascb.org/dora](http://www.ascb.org/dora)). Entries also depend on inclusion in national databases, such as the Chinese Science Citation Database and the Chinese Social Sciences Citation Index. This could open the door to corruption as Chinese universities that publish journals vie for database entries.

China's Discipline Ranking (CDR) system intends to use the list to assess a university's performance by the number of its academics that publish in these journals. Rewards to scientists publishing in the ‘top’ Chinese journals might include payments and questionable promotions, for example, weakening the already distorted evaluation system and impeding the development of science in China.

I suggest that the CDGDC needs to be more service-minded, recognizing that this contentious policy falls outside its authority. Making evaluation systems that are politically independent, non-profit and professional would help to break the CDGDC and CDR monopoly. Universities, too, should rethink the merit of political ranking lists.

**Lihua Yang** *Beihang University, Beijing, China.*  
[lihua.yang@buaa.edu.cn](mailto:lihua.yang@buaa.edu.cn)

## Australia too casual with protection law

The Australian government has set a dangerous precedent in granting exemptions “in the national interest” from its 1999 Environment Protection and Biodiversity Conservation Act. This laxity adversely affects threatened species and disregards scientific advice.

Australia's former environment minister Greg Hunt granted the latest exemption. It permits habitat clearance in Batemans Bay, New South Wales, until the end of this year and dispersal by

“non-lethal means” of camps of the threatened grey-headed flying fox (*Pteropus poliocephalus*). This species is crucial for native forest regeneration. Forced dispersals using loud noise, smoke and lights “often lead to flying fox stress, injuries, or fatalities” (see [go.nature.com/2f9azyl](http://go.nature.com/2f9azyl)) and have proved costly and inefficient (see B. J. Roberts *et al.* *PLoS ONE* **7**, e42532; 2012).

Hunt also signed an exemption in 2014 to permit a shark cull in Western Australia (see [go.nature.com/2f4mrdy](http://go.nature.com/2f4mrdy)), which included the protected great white shark (*Carcharodon carcharias*). Then there is the ongoing destruction of endangered forests in New South Wales, habitats for many threatened species.

In our view, Australia's federal and state governments should be more protective of biodiversity.

**Christian Vincenot** *Kyoto University, Japan.*  
**Sophie Petit** *University of South Australia, Adelaide.*  
[vincenot@i.kyoto-u.ac.jp](mailto:vincenot@i.kyoto-u.ac.jp)

## Renewables from the bottom up

Alan Bernstein and colleagues recommend a global grand-challenge strategy for prioritizing clean-energy research (*Nature* **538**, 30; 2016). In my view, this ‘cockpit’ approach — relying on top-down steering to seek out global solutions — risks galvanizing research along a fixed path.

A diverse set of research initiatives implemented across different scales could be a better approach. That would deliver ways of using energy sustainably under different social and cultural circumstances. The long-term questions for energy research are as much about the evolution of energy use as the revolution in its production.

**Jari Lyytimäki** *Finnish Environment Institute, Helsinki, Finland.*  
[jari.lyytimaki@ymparisto.fi](mailto:jari.lyytimaki@ymparisto.fi)



## NANOSCIENCE

## Single-molecule instant replay

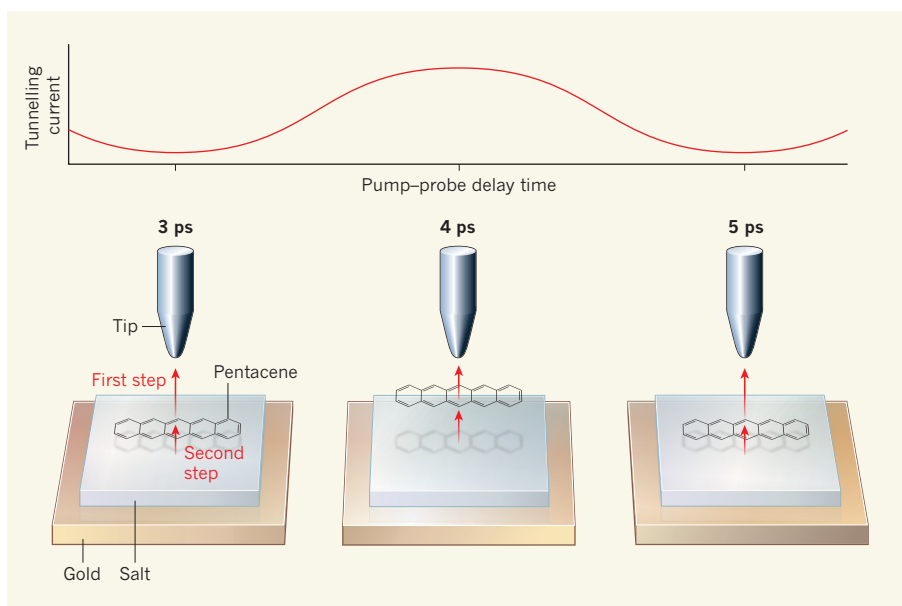
A nanoscale imaging method that uses ultrashort light pulses to initiate and follow the motion of a single molecule adsorbed on a solid surface opens a window onto the physical and chemical dynamics of molecules on surfaces. [SEE LETTER P.263](#)

NICHOLAS CAMILLONE III

Sports enthusiasts know the value of instant replay: footage taken by a high-frame-rate camera, played back at slower speeds, reveals details not obvious to the unaided eye. Tracking a single athlete or a specific body part in this way enables fans to more fully appreciate the action, referees to make crucial rulings, and players and coaches to make performance-enhancing adjustments. On page 263, Cocker *et al.*<sup>1</sup> report a major step towards a single-molecule version of instant replay. They have developed a method for imaging molecules that uses laser pulses to transiently manipulate the electric field at the 'focal point' of a scanning tunnelling microscope. The authors use this technique to track molecular vibrations in real time.

Molecular motions occur in the femtosecond to picosecond regime (1 fs is  $10^{-15}$  s, 1 ps is  $10^{-12}$  s), and involve displacements of the order of picometres. Developing a method that has the spatial and temporal resolution needed to track such motions is a grand challenge, particularly for those studying the chemical physics of molecules that interact with solid surfaces<sup>2,3</sup>. This is because the outcome and efficiency of chemical reactions at surfaces can be linked to features such as single missing surface atoms or steps at the surface that are a single atom high. Furthermore, the chemical bonds in such reactions are broken and formed on picosecond timescales, and the electron and energy transfers that drive the chemistry can happen on femtosecond timescales.

Almost since their invention<sup>4</sup>, scanning tunnelling microscopes (STMs) have made it possible to image single atoms and molecules on solid surfaces<sup>5–7</sup>. STMs consist of an atomically sharp tip held a few tenths of a nanometre from the surface under investigation. When an electrical voltage is applied between the surface and the tip, the quantum-mechanical nature of electrons enables them to 'tunnel' through the energy barrier at the tip–surface gap that, in the absence of actual physical contact, classical mechanics would forbid them to cross. The resulting tunnelling current is exquisitely sensitive to the gap width, and can therefore be used to map the 3D topography of the surface on the picometre scale. This allows



**Figure 1 | Tracking single-molecule oscillations.** Cocker *et al.*<sup>1</sup> placed a pentacene molecule atop a salt film on a gold surface, and held the tip of a scanning tunnelling microscope at a fixed distance above the molecule. The authors excited the system with an ultrashort 'pump' laser pulse, and then induced electrons to tunnel from the gold surface to the tip using a subsequent 'probe' pulse (pulses not shown). Tunnelling occurs in two steps (red arrows): first, an electron tunnels out of the pentacene and into the tip; then another electron tunnels from the gold surface, through the salt film and into the molecule. The authors measured the tunnelling current induced by the probe pulse at different times after the pump pulse (times shown in picoseconds), and observed a sinusoidal modulation with a frequency of approximately 0.5 terahertz. They attribute this modulation to the oscillation of the molecule up and down with respect to the surface.

STMs to 'take a picture' of single molecules on surfaces.

But following a molecule's movements on picosecond timescales has remained an elusive goal. STMs take milliseconds to seconds to take a molecule's picture — too long to capture ultrafast molecular motion directly. A stroboscopic approach is thus desirable, in which an ultrashort pulse of light initiates (pumps) a molecular event at a well-defined point in time, and then, after a well-defined delay, a second pulse (the probe) drives tunnelling by which the molecule is imaged. By measuring the tunnelling current for a fixed pump–probe delay with the STM tip parked at different positions (pixels) across the surface of the molecule, an image representing a snapshot of the molecule at a specific time after the pump can be built up pixel by pixel. Capturing many

such images for a range of pump–probe delays could produce a series of frames constituting a single-molecule instant-replay 'video'.

Cocker *et al.* built on their earlier work<sup>8</sup> to develop a probe scheme that involves a specially engineered support surface, a judiciously chosen molecule and excitation with ultrashort laser pulses. The support consists of an electrically insulating film of salt atop a gold surface (Fig. 1). The molecule (pentacene) has a discrete energy level, occupied by two electrons, that lies at an energy well below the continuum of unoccupied levels in both the gold support and the tungsten STM tip used in the experiments. The pentacene's energy level serves as a bridge for electron transport across the molecule that can be opened or closed.

When a small voltage is applied between the gold and the tip, the bridge remains

closed because the energy level is too low to be involved in electron transport. Under these conditions, tunnelling requires a single, long — and therefore improbable — jump through the entire ‘forbidden’ zone (the salt, the molecule and the vacuum gap). Thus, the tunnelling current is small. But application of a voltage large enough to bring the tip’s unoccupied levels into alignment with the pentacene’s level opens the bridge. This enables a more probable, two-step conduction pathway: electrons can first tunnel out of the pentacene and into the tip, leaving an unoccupied ‘hole’ on the molecule into which an electron from the gold can tunnel through the salt.

Remarkably, the authors found that a pulse of terahertz laser radiation also opens the pentacene bridge, allowing two-step tunnelling in the absence of the electrical voltage. They propose that the pulse’s electric field — which is strongly enhanced at the sharp microscope tip — acts as a transient voltage, switching on the two-step pathway for approximately 100 fs. This represents a completely new mode of single-molecule imaging that enables a snapshot to be captured at a well-defined instant in time.

The characteristics of the terahertz laser pulse are key to the experiment’s success. In simplified terms, the pulse can be thought of as half a wave of electromagnetic radiation. The wave therefore biases the system in only one direction, so that current flows in just one direction. By contrast, a full wave would bias the STM system first in one direction, then in the opposite direction, so that the net current would be almost zero — no image would be taken.

Furthermore, the high reflectivity of metals at THz frequencies (typically greater than 99%)<sup>9,10</sup> minimizes light absorption that would result in heating and thermal expansion of the STM tip. Such expansion can easily result in spurious signals<sup>11</sup> because of the tunnelling current’s extreme sensitivity to the gap width, and has been a major hindrance to the development of laser-excited STM methods.

Cocker *et al.* applied their imaging method to follow molecular motion on picosecond timescales. To do this, they took advantage of the fact that driving an electron out of the occupied level of pentacene using a pump pulse excites vibration of the molecule up and down with respect to the surface. The authors held the tip at a fixed point above the molecule, and used a probe pulse to excite two-step tunnelling at a well-defined time after the pump. Because the tunnelling current is so sensitive to the distance between the molecule and the tip, they could follow the vibrational oscillations of a single molecule in time (Fig. 1), directly measuring not only its frequency (which has previously been measured for single molecules using another STM method<sup>12</sup>), but also its amplitude and phase.

Several imaging methods that have high spatio-temporal resolution are being developed and show promise (for examples, see refs 3, 13–20), but Cocker and colleagues’ approach is unique in its ability to track molecular motion. Nevertheless, further work is needed to develop even more exciting capabilities. For example, a quantitative understanding of the laser-excited tunnelling should be pursued, including detailed computer simulations of the transient electric field generated at the tip by the THz pulse, and of the response of the tip, surface and molecule to that field. It also remains to be seen how this imaging approach may be extended to investigate surface physics and molecular chemistry. Can the damping rates of molecular vibrations be measured? And can molecules be observed as they react on catalytic surfaces?

Single-molecule instant replay will reveal details that will allow dedicated fans — surface physicists and chemists — to appreciate molecular action more deeply. The longer-term challenge is to develop this tool so that referees and coaches — materials scientists and engineers — can use its nanoscale dynamical insights to design, diagnose and improve molecule–surface systems for catalysis, chemical sensors and molecular electronics, all at the single-molecule level. ■

Nicholas Camillone III is in the Chemistry Division, Brookhaven National Laboratory,

Upton, New York 11973-5000, USA.  
e-mail: nicholas@bnl.gov

1. Cocker, T. L., Peller, D., Yu, P., Repp, J. & Huber, R. *Nature* **539**, 263–267 (2016).
2. Lienau, C., Raschke, M. & Ropers, C. in *Attosecond Nanophysics* (eds Hommelhof, P. & Kling, M.) 281–324 (Wiley-VCH, 2014).
3. Petek, H. *ACS Nano* **8**, 5–13 (2014).
4. Binnig, G., Rohrer, H., Gerber, C. & Weibel, E. *Phys. Rev. Lett.* **49**, 57–61 (1982).
5. Baró, A. M. *et al.* *Phys. Rev. Lett.* **52**, 1304–1307 (1984).
6. Chiang, S. *Chem. Rev.* **97**, 1083–1096 (1997).
7. Hansma, P. K. & Tersoff, J. *J. Appl. Phys.* **61**, R1–R23 (1987).
8. Cocker, T. L. *et al.* *Nature Photon.* **7**, 620–625 (2013).
9. Ordal, M. A., Bell, R. J., Alexander, R. W., Newquist, L. A. & Querry, M. R. *Appl. Opt.* **27**, 1203–1209 (1988).
10. Ordal, M. A., Bell, R. J., Alexander, R. W., Long, L. L. & Querry, M. R. *Appl. Opt.* **26**, 744–752 (1987).
11. Gerstner, V., Knoll, A., Pfeiffer, W., Thon, A. & Gerber, G. *J. Appl. Phys.* **88**, 4851–4859 (2000).
12. Stipe, B. C., Rezaei, M. A. & Ho, W. *Science* **280**, 1732–1735 (1998).
13. Terada, Y., Yoshida, S., Takeuchi, O. & Shigekawa, H. *Nature Photon.* **4**, 869–874 (2010).
14. Wu, S. W. & Ho, W. *Phys. Rev. B* **82**, 085444 (2010).
15. Furusawa, K., Hayazawa, N., Catalan, F. C., Okamoto, T. & Kawata, S. *J. Raman Spectrosc.* **43**, 656–661 (2012).
16. Dey, S., Mirell, D., Perez, A. R., Lee, J. & Apkarian, V. A. *J. Chem. Phys.* **138**, 154202 (2013).
17. Lee, J., Perdue, S. M., Rodriguez Perez, A. & Apkarian, V. A. *ACS Nano* **8**, 54–63 (2014).
18. Jahng, J. *et al.* *Appl. Phys. Lett.* **106**, 083113 (2015).
19. Nishiyama, Y., Imaeda, K., Imura, K. & Okamoto, H. *J. Phys. Chem. C* **119**, 16215–16222 (2015).
20. Müller, M., Kravtsov, V., Paarmann, A., Raschke, M. B. & Ernstorfer, R. *ACS Photon.* **3**, 611–619 (2016).

## EVOLUTION

# Genomic remodelling in the primate brain

In many mammals, the gene *Ostn* is expressed in muscles and bones. The discovery that the primate *OSTN* gene has been repurposed to also act in neurons provides clues to how humans evolved their cognitive abilities. [SEE ARTICLE P.242](#)

JUSTINE KUPFERMAN & FRANCK POLLEUX

The cognitive abilities that separate humans and our primate relatives from other mammals are the product of millions of years of evolution, and stem from differences in how our brains develop and function<sup>1</sup>. Although brain maturation in all mammals relies in part on experience-driven development of neuronal circuits, human cognition depends particularly heavily on the experiential learning that occurs during our prolonged period of growth, which lasts up to two decades after birth<sup>2</sup>. The structural and functional changes that shape neuronal circuits during this developmental period are mediated by genes whose transcription is

regulated by neuronal activity<sup>3</sup>. On page 242, Ataman *et al.*<sup>4</sup> describe an unbiased screen to identify genes activated by neuronal excitation in human and mouse neurons. They identify a gene expressed in the bones and muscles of mice and other mammals that, over the course of evolution, was repurposed to act in the neurons of primates.

Because of the importance of experiential learning in humans, identifying gene-expression changes induced by neuronal activity is particularly relevant for understanding the genetic basis of our species’ brain evolution. With this in mind, Ataman *et al.* cultured human neurons *in vitro*. These cultures contain a mix of differentiated cell types found in the brain, including glial cells, which



closed because the energy level is too low to be involved in electron transport. Under these conditions, tunnelling requires a single, long — and therefore improbable — jump through the entire ‘forbidden’ zone (the salt, the molecule and the vacuum gap). Thus, the tunnelling current is small. But application of a voltage large enough to bring the tip’s unoccupied levels into alignment with the pentacene’s level opens the bridge. This enables a more probable, two-step conduction pathway: electrons can first tunnel out of the pentacene and into the tip, leaving an unoccupied ‘hole’ on the molecule into which an electron from the gold can tunnel through the salt.

Remarkably, the authors found that a pulse of terahertz laser radiation also opens the pentacene bridge, allowing two-step tunnelling in the absence of the electrical voltage. They propose that the pulse’s electric field — which is strongly enhanced at the sharp microscope tip — acts as a transient voltage, switching on the two-step pathway for approximately 100 fs. This represents a completely new mode of single-molecule imaging that enables a snapshot to be captured at a well-defined instant in time.

The characteristics of the terahertz laser pulse are key to the experiment’s success. In simplified terms, the pulse can be thought of as half a wave of electromagnetic radiation. The wave therefore biases the system in only one direction, so that current flows in just one direction. By contrast, a full wave would bias the STM system first in one direction, then in the opposite direction, so that the net current would be almost zero — no image would be taken.

Furthermore, the high reflectivity of metals at THz frequencies (typically greater than 99%)<sup>9,10</sup> minimizes light absorption that would result in heating and thermal expansion of the STM tip. Such expansion can easily result in spurious signals<sup>11</sup> because of the tunnelling current’s extreme sensitivity to the gap width, and has been a major hindrance to the development of laser-excited STM methods.

Cocker *et al.* applied their imaging method to follow molecular motion on picosecond timescales. To do this, they took advantage of the fact that driving an electron out of the occupied level of pentacene using a pump pulse excites vibration of the molecule up and down with respect to the surface. The authors held the tip at a fixed point above the molecule, and used a probe pulse to excite two-step tunnelling at a well-defined time after the pump. Because the tunnelling current is so sensitive to the distance between the molecule and the tip, they could follow the vibrational oscillations of a single molecule in time (Fig. 1), directly measuring not only its frequency (which has previously been measured for single molecules using another STM method<sup>12</sup>), but also its amplitude and phase.

Several imaging methods that have high spatio-temporal resolution are being developed and show promise (for examples, see refs 3, 13–20), but Cocker and colleagues’ approach is unique in its ability to track molecular motion. Nevertheless, further work is needed to develop even more exciting capabilities. For example, a quantitative understanding of the laser-excited tunnelling should be pursued, including detailed computer simulations of the transient electric field generated at the tip by the THz pulse, and of the response of the tip, surface and molecule to that field. It also remains to be seen how this imaging approach may be extended to investigate surface physics and molecular chemistry. Can the damping rates of molecular vibrations be measured? And can molecules be observed as they react on catalytic surfaces?

Single-molecule instant replay will reveal details that will allow dedicated fans — surface physicists and chemists — to appreciate molecular action more deeply. The longer-term challenge is to develop this tool so that referees and coaches — materials scientists and engineers — can use its nanoscale dynamical insights to design, diagnose and improve molecule–surface systems for catalysis, chemical sensors and molecular electronics, all at the single-molecule level. ■

Nicholas Camillone III is in the Chemistry Division, Brookhaven National Laboratory,

Upton, New York 11973-5000, USA.  
e-mail: nicholas@bnl.gov

1. Cocker, T. L., Peller, D., Yu, P., Repp, J. & Huber, R. *Nature* **539**, 263–267 (2016).
2. Lienau, C., Raschke, M. & Ropers, C. in *Attosecond Nanophysics* (eds Hommelhof, P. & Kling, M.) 281–324 (Wiley-VCH, 2014).
3. Petek, H. *ACS Nano* **8**, 5–13 (2014).
4. Binnig, G., Rohrer, H., Gerber, C. & Weibel, E. *Phys. Rev. Lett.* **49**, 57–61 (1982).
5. Baró, A. M. *et al.* *Phys. Rev. Lett.* **52**, 1304–1307 (1984).
6. Chiang, S. *Chem. Rev.* **97**, 1083–1096 (1997).
7. Hansma, P. K. & Tersoff, J. *J. Appl. Phys.* **61**, R1–R23 (1987).
8. Cocker, T. L. *et al.* *Nature Photon.* **7**, 620–625 (2013).
9. Ordal, M. A., Bell, R. J., Alexander, R. W., Newquist, L. A. & Querry, M. R. *Appl. Opt.* **27**, 1203–1209 (1988).
10. Ordal, M. A., Bell, R. J., Alexander, R. W., Long, L. L. & Querry, M. R. *Appl. Opt.* **26**, 744–752 (1987).
11. Gerstner, V., Knoll, A., Pfeiffer, W., Thon, A. & Gerber, G. *J. Appl. Phys.* **88**, 4851–4859 (2000).
12. Stipe, B. C., Rezaei, M. A. & Ho, W. *Science* **280**, 1732–1735 (1998).
13. Terada, Y., Yoshida, S., Takeuchi, O. & Shigekawa, H. *Nature Photon.* **4**, 869–874 (2010).
14. Wu, S. W. & Ho, W. *Phys. Rev. B* **82**, 085444 (2010).
15. Furusawa, K., Hayazawa, N., Catalan, F. C., Okamoto, T. & Kawata, S. *J. Raman Spectrosc.* **43**, 656–661 (2012).
16. Dey, S., Mirell, D., Perez, A. R., Lee, J. & Apkarian, V. A. *J. Chem. Phys.* **138**, 154202 (2013).
17. Lee, J., Perdue, S. M., Rodriguez Perez, A. & Apkarian, V. A. *ACS Nano* **8**, 54–63 (2014).
18. Jahng, J. *et al.* *Appl. Phys. Lett.* **106**, 083113 (2015).
19. Nishiyama, Y., Imaeda, K., Imura, K. & Okamoto, H. *J. Phys. Chem. C* **119**, 16215–16222 (2015).
20. Müller, M., Kravtsov, V., Paarmann, A., Raschke, M. B. & Ernstorfer, R. *ACS Photon.* **3**, 611–619 (2016).

## EVOLUTION

# Genomic remodelling in the primate brain

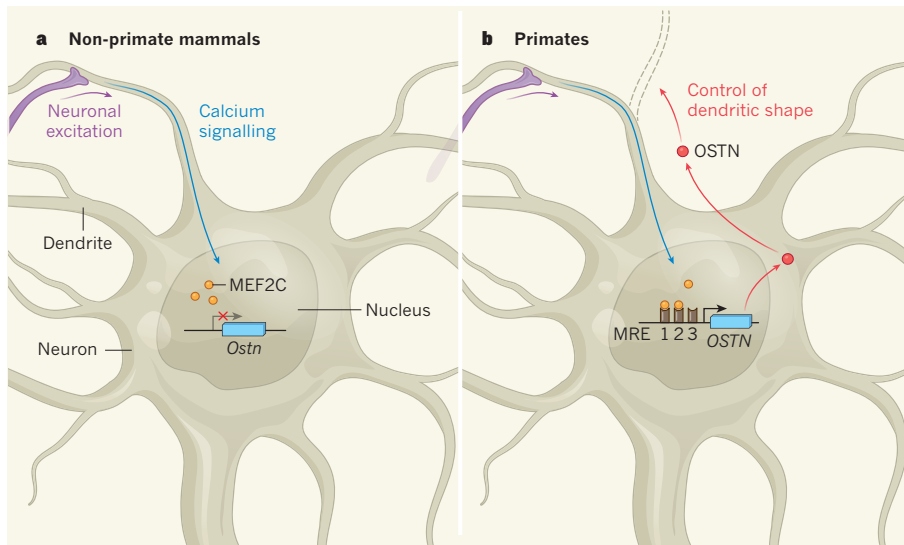
In many mammals, the gene *Ostn* is expressed in muscles and bones. The discovery that the primate *OSTN* gene has been repurposed to also act in neurons provides clues to how humans evolved their cognitive abilities. [SEE ARTICLE P.242](#)

JUSTINE KUPFERMAN & FRANCK POLLEUX

The cognitive abilities that separate humans and our primate relatives from other mammals are the product of millions of years of evolution, and stem from differences in how our brains develop and function<sup>1</sup>. Although brain maturation in all mammals relies in part on experience-driven development of neuronal circuits, human cognition depends particularly heavily on the experiential learning that occurs during our prolonged period of growth, which lasts up to two decades after birth<sup>2</sup>. The structural and functional changes that shape neuronal circuits during this developmental period are mediated by genes whose transcription is

regulated by neuronal activity<sup>3</sup>. On page 242, Ataman *et al.*<sup>4</sup> describe an unbiased screen to identify genes activated by neuronal excitation in human and mouse neurons. They identify a gene expressed in the bones and muscles of mice and other mammals that, over the course of evolution, was repurposed to act in the neurons of primates.

Because of the importance of experiential learning in humans, identifying gene-expression changes induced by neuronal activity is particularly relevant for understanding the genetic basis of our species’ brain evolution. With this in mind, Ataman *et al.* cultured human neurons *in vitro*. These cultures contain a mix of differentiated cell types found in the brain, including glial cells, which



**Figure 1 | A gene repurposed.** Neurotransmission between neurons leads to intracellular calcium signalling, which in turn activates the transcription factor MEF2C. This protein binds to DNA sequences called MEF2-responsive elements (MREs) to induce transcription of nearby genes. **a**, Ataman *et al.*<sup>4</sup> report that non-primate mammals do not have MEF2C binding sites in the regulatory sequences that lie upstream of the osteonin (*Ostin*) gene, and their neurons do not express *Ostin*. **b**, By contrast, three MEF2C sites have evolved in this region in primates, and so the equivalent primate gene, *OSTN*, is transcribed in response to neuronal activity (it is not clear whether MRE3 is required for activity-dependent *OSTN* expression). *OSTN* protein is secreted from neurons and can regulate branching of neuronal projections called dendrites, perhaps regulating the structural changes that neurons undergo during learning in primates.

support neurons and promote formation of the synaptic connections between them, and several subtypes of neuron from the brain's cortex. The authors stimulated these cultures to mimic increased neuronal activity, thereby increasing calcium signalling within neurons and inducing activity-dependent gene transcription.

Next, Ataman and colleagues used RNA sequencing to identify upregulated transcripts and confirmed that several well-characterized<sup>5,6</sup> 'immediate early' genes were rapidly induced within one hour of increased neuronal activity, including the transcription factors *NPAS4* and *FOSB*. They also identified a set of late-response genes induced within six hours, which included previously reported genes<sup>7</sup> such as *BDNF*. At this late time point, the authors found a few transcripts specific to human neurons. The most highly induced of these genes was that for osteonin (*OSTN*).

In mice, the corresponding gene *Ostin* encodes a secreted protein that is involved in glucose metabolism in muscles and bones<sup>8,9</sup>. *Ostin* is not expressed in the mouse brain, and Ataman *et al.* found that its expression could not be induced by neuronal activity in mouse neuronal cultures. Therefore, *OSTN* expression seems to be regulated by neuronal activity in human but not mouse neurons — a supposition that the authors corroborated through several lines of investigation.

Ataman and colleagues reported that *OSTN* was widely expressed throughout the human neocortex (which is involved in

higher cognitive functions, including sensory processing) and was highly enriched in the mature neurons of the developing cortex. They also found that, in the primary visual cortex of macaques, *OSTN* expression was induced by increased sensory-evoked neuronal activity. Together, these lines of evidence point to activity-dependent expression of *OSTN* in primate neurons *in vivo*.

At the genetic level, what underlies this shift from *Ostin* transcription in the bones of mice to activity-dependent regulation of *OSTN* transcription in primate neurons? To address this question, the authors first demonstrated

**In primates but not in other mammals, *OSTN* might regulate structural changes that neurons undergo during learning.**

that *OSTN* expression in human neurons is regulated by a 2-kilobase-long promoter region that lies immediately upstream gene. By engineering a series of truncations and point mutations in this region, the researchers identified a minimal 85-base-long region that mediates gene activation in response to neuronal activity. In primates, this region contains three short DNA sequences to which transcription factors of the myocyte enhancer factor 2 (MEF2) family, such as MEF2C, can bind. In mice, however, the *Ostin* promoter does not contain these MEF2-responsive elements (MREs).

MEF2 transcription factors are crucial for

activity-dependent transcription in neurons and have a role in key aspects of neuronal development<sup>10,11</sup>. The researchers found that two of the three MREs are highly evolutionarily conserved between humans and anthropoid primates, but that there are several differences in these sequences in prosimian primates, rodents, dolphins and several other mammalian species that result in the absence of MREs in the *Ostin* promoter of these species. By replacing the three primate MREs with the equivalent mouse sequences and monitoring gene transcription, the authors convincingly demonstrated that the evolutionary switch to activity-dependent transcription of the *OSTN* gene in primates emerged through a few single-nucleotide mutations that created MREs (Fig. 1).

What is the function of *OSTN* in primate neurons? The authors overexpressed or repressed *OSTN* in human neuronal cultures, and discovered that expression of the gene regulates the shape of dendrites — the branched parts of neurons that receive and integrate synaptic information from other neurons. This result suggests that, in primates but not in other mammals, *OSTN* might regulate structural changes that neurons undergo during learning. Indeed, *OSTN* belongs to a family of genes that encode secreted protein fragments called natriuretic peptides, which have been shown<sup>12</sup> to promote branching of neuronal projections called axons.

Further experiments will be needed to fully determine the impact of *OSTN* expression on primate brain development. Defining the roles of *OSTN* — and of the other activity-dependent genes identified in Ataman and colleagues' study — will improve our understanding of the evolutionary mechanisms that enabled the emergence of primate-specific features of brain development and function. ■

**Justine Kupferman and Franck Polleux** are in the Department of Neuroscience, the Mortimer B. Zuckerman Mind Brain Behavior Institute and the Kavli Institute for Brain Science, Columbia University, New York, New York 10032, USA.  
e-mails: [jb2524@cumc.columbia.edu](mailto:jb2524@cumc.columbia.edu); [fp2304@cumc.columbia.edu](mailto:fp2304@cumc.columbia.edu)

1. Carroll, S. B. *Nature* **422**, 849–857 (2003).
2. Petanjek, Z. *et al. Proc. Natl Acad. Sci. USA* **108**, 13281–13286 (2011).
3. Greer, P. L. & Greenberg, M. E. *Neuron* **59**, 846–860, (2008).
4. Ataman, B. *et al. Nature* **539**, 242–247 (2016).
5. Kim, T. K. *et al. Nature* **465**, 182–187 (2010).
6. Lanahan, A. & Worley, P. *Neurobiol. Learn. Mem.* **70**, 37–43 (1998).
7. Hughes, P., Beilharz, E., Gluckman, P. & Dragunow, M. *Neuroscience* **57**, 319–328 (1993).
8. Moffatt, P. *et al. J. Biol. Chem.* **282**, 36454–36462 (2007).
9. Subbotina, E. *et al. Proc. Natl Acad. Sci. USA* **112**, 16042–16047 (2015).
10. Flavell, S. W. *et al. Science* **311**, 1008–1012 (2006).
11. Mao, Z., Bonni, A., Xia, F., Nadal-Vicens, M. & Greenberg, M. E. *Science* **286**, 785–790 (1999).
12. Zhao, Z. & Ma, L. *Proc. Natl Acad. Sci. USA* **106**, 18016–18021 (2009).



## GEOSCIENCE

# Cliffs make a hasty retreat

Knowing how quickly coastlines have eroded is essential for quantifying the future risks to both infrastructure and coastal ecosystems. Writing in *Proceedings of the National Academy of Sciences*, Hurst *et al.* report retreat rates for cliffs on the south coast of Britain (pictured) over thousands of years, and show that this erosion has speeded up in recent centuries (M. D. Hurst *et al. Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1613044113>; 2016).

The authors combine high-precision

isotopic dating with a detailed numerical model of coastal erosion. They find that retreat rates were relatively slow (2–6 centimetres per year) until a few hundred years ago before subsequently accelerating (reaching 22–32 cm yr<sup>-1</sup>), and suggest that this is driven by a combination of environmental and anthropogenic factors. The results show that a careful analysis of long-term coastal records can uncover more-recent changes, which will help in developing predictive models. **Ryan Wilkinson**



ANDREW THOMAS/GETTY

## CANCER

# Bad neighbours cause bad blood

**Expression of a blood–cancer–associated genetic mutation in the non–blood cells of the bone marrow is sufficient to cause blood cancer in mice. This finding could point to new approaches to treating an often–fatal disease. [SEE LETTER P.304](#)**

GORDON CHAN & BENJAMIN G. NEEL

**M**utations in blood stem cells can result in excessive production of one or more types of mature blood cell and a form of cancer called a myeloproliferative neoplasm. One such neoplasm is juvenile myelomonocytic leukaemia (JMML), a rare childhood disease characterized by expansion of the white blood cell population<sup>1</sup>. Currently, JMML can be treated only by stem-cell transplantation, but about 50% of patients relapse following this procedure. A more detailed

understanding of the disease could lead to new therapies and improved outcomes. Dong *et al.*<sup>2</sup> report on page 304 that mutation of the gene *Ptpn11* in the non-blood cells that surround blood stem cells causes a JMML-like cancer in mice, owing to altered interactions between these two cell types.

Mutations in genes that encode proteins of the RAS–ERK signalling pathway are found in the blood cells (but no other cells) of about 90% of children with sporadic JMML (refs 1, 3). The most frequent of these occur in the gene *PTPN11* (refs 1, 3). Such mutations activate



## 50 Years Ago

Recent reports on the chromosomes and total DNA of the chicken cell nucleus show that they differ from those of typical placental mammals in an unusual pattern of chromosome sizes and a DNA content equal to half of that of most mammals ... anomalies include incomplete dominance ... and alteration of phenotypic sex without visible alterations in the karyotype. One explanation ... could be that three-fourths of the normal genetic dose occurs fairly frequently in the chicken. Three-fourths of the normal gene dose could occur if each of the arms of the major chromosomes in the chicken were isochromosomes ... which double back on themselves for transcription but straighten out and are apposed in the orthodox manner for meiosis ... A chromosome pair would provide four readings of the message, and a cell with three out of four of the normal readings might well be viable ... One evolutionary advantage of the looped isochromosome system proposed here might be that such a chromosome would be better able to withstand the damage to the DNA ... caused by the relatively high internal temperature of birds. **From *Nature* 12 November 1966**

## 100 Years Ago

In *Nature* of October 5 it is stated that the cause of the optical deterioration of the atmosphere in July and August “is for the time being still in doubt. Up to the present no reports of volcanic eruptions have come to hand from any part of the globe.” I wish to state that in July a strong outburst occurred of the Stromboli volcano ... there was an extraordinary eruption of fluid and incandescent lava to a great height, followed by a thick rain of lapilli and ashes. **From *Nature* 9 November 1916**

## CANCER

# Bad neighbours cause bad blood

**Expression of a blood–cancer–associated genetic mutation in the non–blood cells of the bone marrow is sufficient to cause blood cancer in mice. This finding could point to new approaches to treating an often–fatal disease.**

GORDON CHAN & BENJAMIN G. NEEL

Mutations in blood stem cells can result in excessive production of one or more types of mature blood cell and a form of cancer called a myeloproliferative neoplasm. One such neoplasm is juvenile myelomonocytic leukaemia (JMML), a rare childhood disease characterized by expansion of the white blood cell population<sup>1</sup>. Currently, JMML can be treated only by stem-cell transplantation, but about 50% of patients relapse following this procedure. A more detailed understanding of the disease could lead to new therapies and improved outcomes. In a paper online in *Nature*, Dong *et al.*<sup>2</sup> report that mutation of the gene *Ptpn11* in the non-blood cells that surround blood stem cells causes a JMML-like cancer in mice, owing to altered interactions between these two cell types.

Mutations in genes that encode proteins of the RAS–ERK signalling pathway are found in the blood cells (but no other cells) of about 90% of children with sporadic JMML (refs 1, 3). The most frequent of these occur in the gene *PTPN11* (refs 1, 3). Such mutations activate the enzyme encoded by *PTPN11*, called SHP2, increasing RAS–ERK signalling. Certain *PTPN11* mutations can be inherited and so are present in all cells of the body — this leads to the developmental disorder Noonan syndrome<sup>4</sup>. This syndrome is characterized by cardiac, facial, cognitive and growth abnormalities, but about 20% of affected children also have abnormal white blood cell counts, and a few develop a congenital form of JMML.

Previous studies of JMML have focused on the effects of disease-associated *PTPN11* in blood cells, and have established that expression of the mutant gene in these cells alone can promote development of a JMML-like myeloproliferative neoplasm (MPN) in mice<sup>5,6</sup>. Dong *et al.* confirmed that blood-cell-restricted expression of mutant *Ptpn11* causes MPN (Fig. 1a). They next studied mice that harbour *Ptpn11* mutations only in stromal (non-blood) cell types of the bone marrow, which surround

the developing blood cells. Surprisingly, these mice, which have genetically normal blood cells, develop a disease similar to congenital JMML, in which the mutant gene is present in all tissues.

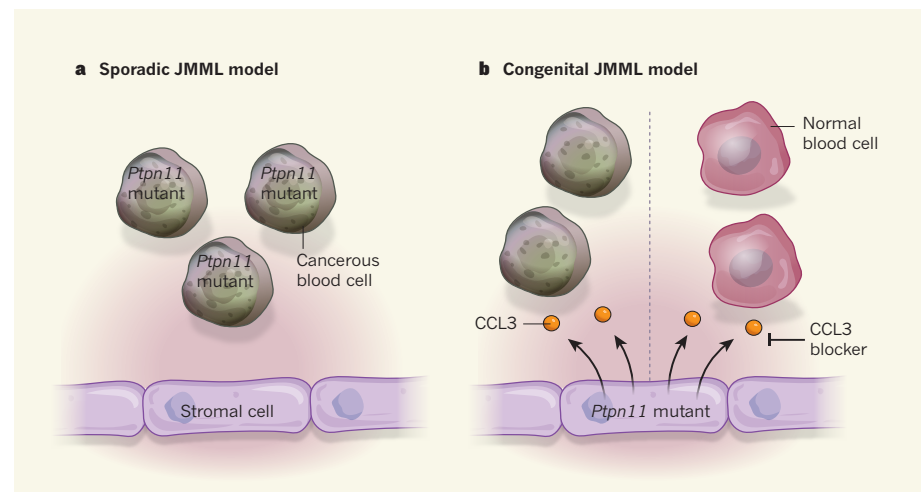
The authors report that mutant *Ptpn11* can give rise to MPN when expressed in some stromal cells of the bone marrow (mesenchymal stem cells and the bone progenitors that they give rise to), but not in others (mature bone or blood-vessel cells). Furthermore, if genetically normal blood stem cells reside long enough in this abnormal microenvironment, they become fully cancerous and can give rise to MPN when transplanted into healthy mice. Presumably, the microenvironment induces genetic mutations or epigenetic changes — modifications that affect gene expression without changing DNA sequence — in normal blood stem

cells that result in this transformation.

What these changes in normal blood stem cells are, and whether they affect genes of the RAS–ERK pathway, is a question left unresolved. Perhaps clarifying this issue would shed light on the causes of the approximately 5–10% of JMML cases for which a genetic basis is undefined<sup>1</sup>. For example, might these cancers be caused by mutations in mesenchymal stem cells, rather than blood stem cells?

The finding that an abnormal bone-marrow microenvironment can promote blood disease is not unprecedented: a variety of genetic alterations can evoke MPN or acute leukaemia when engineered into the bone-marrow stromal cells of mice<sup>7</sup>. But Dong and colleagues model an actual human disease — JMML associated with Noonan syndrome — in which an aberrant bone-marrow microenvironment is present from birth. Their results therefore suggest that such a microenvironment might provide the initial, and potentially the more potent, MPN-promoting stimulus in congenital JMML.

The levels of SHP2 activation caused by *PTPN11* mutation in patients can vary<sup>1,3</sup>. Only the most strongly activating mutations cause JMML in children with Noonan syndrome. Perhaps, under such circumstances, JMML can develop only when both blood cells and the stromal microenvironment bear the mutant gene. A similar requirement has been reported for a mouse model of



**Figure 1 | Collaboration causes blood cancer.** **a**, Mutation in the gene *Ptpn11* in the blood cells of mice and humans causes the cancer juvenile myelomonocytic leukaemia (JMML), which is one of a family of cancers called myeloproliferative neoplasms (MPNs). The stromal cells that surround developing blood cells do not need to be mutated for sporadic JMML to develop. **b**, However, Dong *et al.*<sup>2</sup> report that expression of mutant *Ptpn11* in the stroma causes MPN in mice. These mice provide insight into a congenital form of JMML in which the mutation is present in all cells. The mutant stromal cells produce pro-inflammatory factors, including the protein CCL3, which cause cancerous transformation of normal blood cells into cells of a JMML-like MPN. The incidence of stroma-induced MPNs can be reduced by drugs that block CCL3 action.



MPN caused by another genetic defect<sup>8</sup>.

Dong *et al.* find that several pro-inflammatory factors produced by mutant stromal cells, including the protein CCL3, are present in blood serum from stromal-cell-specific *Pttn11* mutant mice and from people with Noonan syndrome. These findings suggest that such factors could serve as biomarkers to identify children at increased risk of JMML. Furthermore, the authors find that drugs that block CCL3 action reduce the incidence of MPN in mice (Fig. 1b). Administering such drugs to high-risk patients might spare them from requiring stem-cell transplantation.

The current study also has potential implications for sporadic JMML. Previous studies showed that aberrant blood cells can send signals to stromal cells that alter the bone-marrow microenvironment to favour the blood cells' own expansion or to impede the survival, proliferation and function of normal blood stem cells<sup>7</sup>. Dong *et al.* find that, like the mutant stroma, *Pttn11* mutant blood cells produce pro-inflammatory factors. This suggests a vicious circle in which mutant blood cells perturb the microenvironment, which in turn augments expansion of the cancerous blood-cell population. If bone-marrow stromal cells have been rendered abnormal by cancerous

blood cells in sporadic JMML, persistence of this abnormal microenvironment might contribute to disease recurrence after a stem-cell transplant.

Consequently, perhaps the most immediate intervention suggested by Dong and colleagues' work is to investigate the effects of CCL3 blockers in transplant regimens for children with JMML. One such drug, maraviroc, is already used to treat HIV, and also blocks graft-versus-host disease — a complication of bone-marrow transplantation<sup>9</sup>. The authors find that such agents are more effective at preventing MPN post-transplant than in combating established MPN in mice. Nevertheless, combining these blockers with targeted therapies aimed at reversing the effects of increased RAS-ERK activation could represent a promising therapeutic strategy for JMML.

Finally, the results of the current study resonate with, and inform, emerging themes in cancer biology. It is becoming clear that the interactions between cancerous cells and their microenvironment determine cancer behaviour and response to therapy. It is also known that inflammation can initiate, as well as enhance, cancer. In solid tumours, one major effect of tumour-associated inflammation is to suppress the host immune system<sup>10</sup>. It will be

interesting to see whether therapies directed against the JMML microenvironment evoke or augment an anti-MPN immune response. ■

**Gordon Chan** is in the Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario M5S 3M2, Canada.

**Benjamin G. Neel** is at the Laura and Isaac Perlmutter Cancer Center, New York University Langone Medical Center, New York, New York 10016, USA.

e-mails: chan.mail@utoronto.ca;  
benjamin.neel@nyumc.org

1. Chang, T. Y., Dvorak, C. C. & Loh, M. L. *Blood* **124**, 2487–2497 (2014).
2. Dong, L. *et al.* *Nature* <http://dx.doi.org/10.1038/nature20131> (2016).
3. Chan, G. & Neel, B. G. in *Protein Tyrosine Phosphatases in Cancer* (eds Neel, B. G. & Tonks, N. K.) 115–143 (Springer, 2016).
4. Roberts, A. E., Allanson, J. E., Tartaglia, M. & Gelb, B. D. *Lancet* **381**, 333–342 (2013).
5. Mohi, M. G. *et al.* *Cancer Cell* **7**, 179–191 (2005).
6. Xu, D. *et al.* *J. Exp. Med.* **208**, 1977–1988 (2011).
7. Hoggatt, J., Kfoury, Y. & Scadden, D. T. *Annu. Rev. Pathol.* **11**, 555–581 (2016).
8. Walkley, C. R., Shea, J. M., Sims, N. A., Purton, L. E. & Orkin, S. H. *Cell* **129**, 1081–1095 (2007).
9. Woollard, S. M. & Kanmogne, G. D. *Drug Des. Dev. Ther.* **9**, 5447–5468 (2015).
10. Shalapour, S. & Karin, M. *J. Clin. Invest.* **125**, 3347–3355 (2015).

## In Retrospect

## Eighty years of stress

**The discovery in 1936 that rats respond to various damaging stimuli with a general response that involves alarm, resistance and exhaustion launched the discipline of stress research.**

GEORGE FINK

Modern research into biological stress — the physiological response of an organism to stressful stimuli — spans broad disciplines, from genetics to endocrinology to brain imaging. Much of this diverse field has a shared origin in a landmark paper published 80 years ago in *Nature*<sup>1</sup>. In this brief note, the biologist Hans Selye described a common physiological response in rats subjected to a diverse range of harmful factors; he named this the stress response<sup>2</sup>. Selye's observation set the scene for decades of future discoveries.

The history of stress research begins in the mid-nineteenth century with the physiologist Claude Bernard<sup>3</sup>, who proposed that the cells of the body are bathed in a fixed internal environment that is maintained in the face of changing external conditions by compensatory

physiological changes. In the early twentieth century, this concept was termed homeostasis by another physiologist, Walter Cannon<sup>4</sup>. Cannon described how animals react to stimuli that threaten homeostasis with a physiological response that he dubbed 'fight or flight'.

In his 1936 paper, Selye reported that when rats were subjected to nonspecific damaging agents such as exposure to cold, surgical injury or intoxication with diverse drugs, they showed a typical response that was not dependent on the nature of the agent. He described three phases to this general adaptation syndrome (GAS): alarm, resistance and exhaustion. Alarm corresponds to the fight-or-flight response described by Cannon; resistance to a period in which the body adapts to repeated exposure to the stress; and exhaustion to a relapse of symptoms that occurs if the stress is exerted for too long. The main features of the syndrome were suppression of

the immune system, ulceration of the lining of the stomach and small intestine, and activation of the two main neurally activated stress-response systems — the hypothalamic–pituitary–adrenocortical (HPA) axis and the sympathomedullary system (Fig. 1).

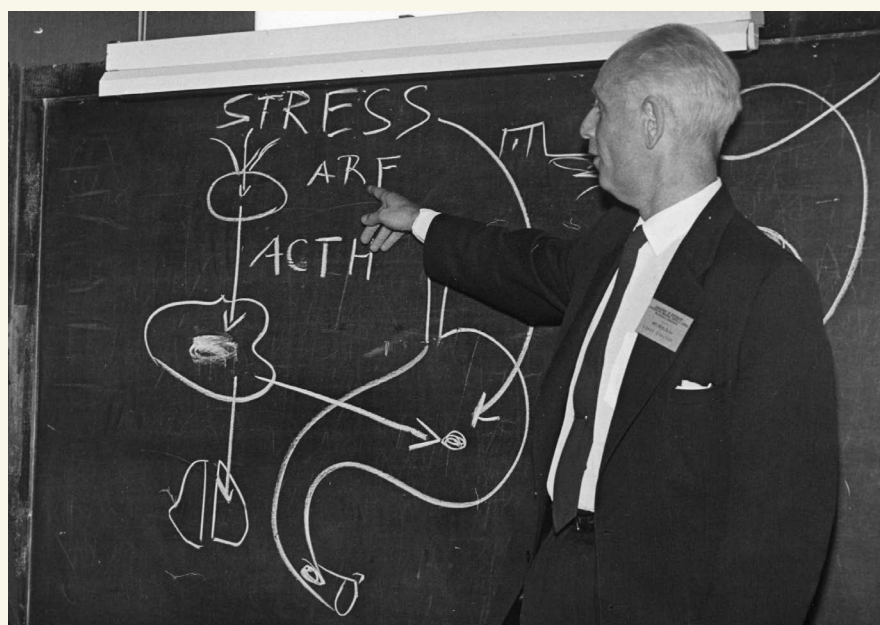
In the HPA axis, activation of nerves in the brain's hypothalamus triggers release of the hormone corticotropin-releasing factor (CRF), which in turn causes secretion of the hormone adrenocorticotropin (ACTH) from the pituitary gland at the base of the brain into the blood. ACTH stimulates secretion of hormones, including cortisol, from the cortex of the adrenal gland. Hypothalamic signalling also triggers the sympathomedullary system, which increases heart rate, shunts blood from the skin and gut towards the skeletal muscles, and releases adrenaline from the medulla region of the adrenal gland. Adrenaline and cortisol act synergistically to increase levels of glucose in the blood<sup>5</sup>, providing the energy required for fight or flight.

Selye defined stress as "the nonspecific response of the body to any demand"<sup>2,6,7</sup>. He believed that stress was different from emotional arousal or nervous tension<sup>6</sup> because, as he wrote, the same general response was known to occur under or in response to anaesthesia in humans and animals, and also to occur in plants and bacteria, which have no nervous system (the generality of this response is reviewed in refs 8 and 9). Thus, he concluded, stress has a key protective role in all organisms. Moreover, he correctly proposed that the stress response is the same whether a stimulus is pleasant or unpleasant — the key factor is whether the intensity of the stimulus necessitates adaptation<sup>7</sup>.

The biologist's views and definition of stress were widely accepted, although not without resistance<sup>6,10</sup>. In fact, recognition of the fact that stress is not necessarily due to nervous arousal encouraged many scientists and clinicians to use the more precise terms neurogenic stress and psychogenic stress<sup>6</sup>.

Selye recognized that maintaining an internal physiological balance through homeostasis (stability through constancy) could not by itself ensure the stability of body systems under stress, and he coined the term heterostasis to describe the process by which a new steady state could be achieved through adaptive mechanisms<sup>7</sup>. The idea of heterostasis can be considered a forerunner to the discovery, in the late 1980s, that neural regulation of feedback in various body systems can alter physiological responses to enable animals to meet a stressful challenge<sup>11–13</sup>: a phenomenon known as allostasis (stability through change).

From allostasis has sprung the idea that the amount of stress to which an individual



**Figure 1 | Hans Selye outlines the general adaptation syndrome.** In 1936, the biologist Hans Selye observed a stress response in rats subjected to a range of potentially harmful stimuli<sup>1</sup>. The syndrome involves three phases: alarm, resistance and exhaustion (ARE). Release of the hormone adrenocorticotropin (ACTH) in response to stress stimulates secretion of multiple hormones from the adrenal gland, affecting many organs and systems, including the thymus glands and stomach (pictured).



arousal encouraged many scientists and clinicians to use the more precise terms neurogenic stress and psychogenic stress<sup>6</sup>.

Selye recognized that maintaining an internal physiological balance through homeostasis (stability through constancy) could not by itself ensure the stability of body systems under stress, and he coined the term heterostasis to describe the process by which a new steady state could be achieved through adaptive mechanisms<sup>7</sup>. The idea of heterostasis can be considered a forerunner to the discovery, in the late 1980s, that neural regulation of feedback in various body systems can alter physiological responses to enable animals to meet a stressful challenge<sup>11–13</sup>: a phenomenon known as allostasis (stability through change).

From allostasis has sprung the idea that the amount of stress to which an individual is subjected over time can be quantified by measuring adverse effects on the cardiovascular and other organ systems<sup>13</sup>. Further research is required to determine how best to quantify this allostatic load to provide a robust index of stress. It also remains to be seen whether and how such an index could be applied in the clinic.

A growing understanding of stress led to an ability to modulate stress-response pathways to treat disease. In 1950, the Nobel Prize in Physiology or Medicine was awarded for the first use of synthetic adrenocortical hormones, such as synthetic cortisol, as powerful anti-inflammatory agents for treating conditions such as rheumatoid arthritis.

In 1981, the amino-acid sequence of CRF was characterized<sup>14</sup>. Selye died in 1982, and so did not witness the quantum leap in our understanding of stress that followed. The structure of CRF, the characterization of its receptors, and the development of investigative tools such as molecular genetics and brain imaging have allowed us to map the structure and function of brain pathways involved in the stress response. These advances have facilitated our understanding of the neural mechanisms involved in many mental disorders, including anxiety, depression and post-traumatic stress, and provided targets for therapies to treat these disorders.

However, many uncertainties remain<sup>15</sup>. These include the extent to which genetic mechanisms determine susceptibility or resilience to stress, and whether some people are genetically more susceptible to post-traumatic stress than others. It is unclear whether genetic predisposition to stress susceptibility can be affected by the environment, either in the womb or after birth, and whether this type of susceptibility can be transmitted between generations. Another question facing researchers is whether stress-induced changes in brain chemistry, which occur in major depression and anxiety, might damage the human brain, reversibly or irreversibly.

The holy grail for many stress neurobiologists is to find a way to selectively extinguish

adverse post-traumatic memories. The chances of such a discovery in the near future are perhaps remote, given the complexity of the human brain and the accompanying ethical issues. Nonetheless, it would doubtless benefit many people who have post-traumatic stress and other mental disorders.

We now know that, although GAS as described by Selye sometimes manifests in people who experience extreme stress such as septic shock or multiple traumatic injuries, some components of GAS, such as stomach ulcers, are not useful quantitative clinical indicators of stress. Modern markers of stress include behavioural observations and hormonal measurements of sympathomedullary and HPA activation. As such, the scientific and clinical currency of GAS has diminished over time. Nonetheless, it retains its importance in launching the stress field. Eighty years later, the foundations laid by Selye are still being built on. ■

George Fink is at the Florey Institute of Neuroscience and Mental Health, University of Melbourne, Parkville,

Victoria 3010, Australia.  
e-mail: george.fink@florey.edu.au

1. Selye, H. *Nature* **138**, 32 (1936).
2. Selye, H. *Br. Med. J.* **1**, 1383–1392 (1950).
3. Bernard, C. *Leçons sur les propriétés physiologiques et les altérations pathologiques des liquides de l'organisme* (Baillière, 1859).
4. Cannon, W. B. *The Wisdom of the Body* (Norton, 1932).
5. Sherwin, R. R. & Sacca, L. *Am. J. Physiol. Endocrinol. Metab.* **245**, E157–E165 (1984).
6. Selye, H. *J. Hum. Stress* **1**, 37–44 (1975).
7. Selye, H. *Stress in Health and Disease* (Butterworth, 1976).
8. Haslbeck, M. & Vierling, E. *J. Mol. Biol.* **427**, 1537–1548 (2015).
9. Park, C.-J. & Seo, Y.-S. *Plant Pathol. J.* **31**, 323–333 (2015).
10. Pacak, K. *et al. Am. J. Physiol.* **275**, R1247–R1255 (1998).
11. Sterling, P. & Eyer, J. in *Handbook of Life Stress, Cognition and Health* (eds Fisher, S. & Reason, J.) 629–649 (Wiley, 1988).
12. Schulkin, J. (ed.) *Allostasis, Homeostasis, and the Costs of Physiological Adaptation* (Cambridge Univ. Press, 2004).
13. McEwen, B. S. *Physiol. Rev.* **87**, 873–904 (2007).
14. Vale, W., Spiess, J., Rivier, C. & Rivier, J. *Science* **213**, 1394–1397 (1981).
15. Fink, G. *J. Neuroendocrinol.* **23**, 107–117 (2011).

This article was published online on 26 October 2016.

#### ATOMIC PHYSICS

## A strange kind of liquid

**Interactions between the magnetic dipoles of dysprosium atoms in an ultracold gas can produce a 'self-bound' droplet. This provides a useful isolated system for probing the quantum-mechanical properties of ultracold gases. SEE LETTER P.259**

BRUNO LABURTHE-TOLRA

**U**ltracold atomic gases are produced at temperatures below 1 microkelvin, and are seven to eight orders of magnitude less dense than typical liquids such as water<sup>1</sup>. However, because the kinetic energy of these gases is dominated by the interactions between their atoms, they can display behaviour akin to that of liquids, and even superfluidity — transport without friction — at temperatures approaching absolute zero. Such systems therefore act both like a gas (they expand freely in the absence of a trap) and also much like superfluid liquid helium. Physicists call them quantum fluids. On page 259, Schmitt *et al.*<sup>2</sup> show that when interactions between magnetic dipoles are in play, quantum fluids display another typical feature of liquids: they form droplets.

In ultracold atomic gases, dipole–dipole interactions can arise because the atoms have tiny, but non-zero, magnetic dipole moments. Usually, the associated dipole–dipole interaction is extremely small relative to the short-range force between atoms — the van der Waals interaction. However, there are a few atoms for which dipole–dipole interactions can

have abnormally large amplitudes. Earlier this year, Schmitt and colleagues' group showed that an ultracold gas of dysprosium atoms, held in a trap produced by a tightly focused laser beam, undergoes a dramatic change if dipole–dipole interactions are large enough relative to short-range interactions<sup>3</sup>. Instead of the gas filling the trap homogeneously, an ordered arrangement of droplets spontaneously appears (Fig. 1a).

The predisposition to form structures is intimately associated with the long-range nature of dipole–dipole interactions. In fact, many systems whose members interact with each other at relatively long distances share a tendency to form self-assembled arrangements. Well-known examples range from astrophysical systems driven by gravitational forces to microscopic systems such as self-organized crystals of ultracold ions formed by electrostatic forces<sup>4</sup>. In the case of dipole–dipole interactions, stable structures are expected to be sensitive to the geometry of the trap because the magnetic fields associated with the dipoles are anisotropic (direction-dependent).

In the current paper, Schmitt *et al.* introduce two subtle modifications to their previous

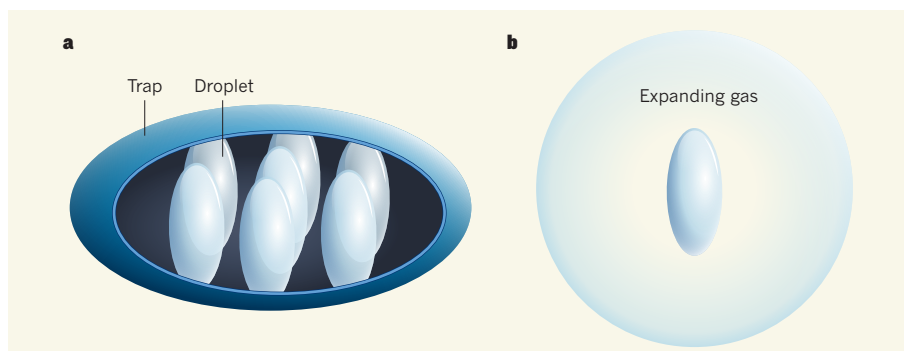
set-up. First, they produce a trap that is elongated in a direction parallel to the magnetic dipoles of the atoms. Instead of an ensemble of droplets, the atomic gas forms single, 'self-bound' droplets (Fig. 1b).

Schmitt and collaborators' second and decisive change is to remove the trap after the droplets have been produced, and to replace it with a simple magnetic-field gradient whose only effect is to compensate for gravity. This can be seen as an inexpensive way to perform the experiment in microgravity. Instead of expanding in space like a normal gas, the droplets remain localized. This is because the dipole-dipole interactions between the atoms provide attraction and cohesion, in a similar way to the hydrogen bonds between H<sub>2</sub>O molecules in water.

Perhaps one of the most interesting aspects of Schmitt and colleagues' work is their demonstration that there are a critical number of atoms below which the liquid droplets are unstable and turn into a gas. This phase transition is one of the first to be observed in a cold-atom experiment without a trap. To understand this transition, it is necessary to explain why the liquid droplets can be stable in the first place. The authors observe self-bound droplets when the balance between repulsive short-range interactions and attractive long-range interactions would have been expected to lead to collapse. However, collapse does not occur because van der Waals and dipole-dipole interactions create quantum fluctuations<sup>5,6</sup>. These fluctuations apply a pressure to the droplets that rapidly increases with density, and is therefore sufficient to ensure stability only if the number of atoms is large enough.

The main novelty of the authors' work is that their extremely dilute droplets are self-bound — a property that is reminiscent of solitonic structures. Solitons were named after non-dispersive, 'solitary' waves of water, but are well known in optics and also exist in the context of quantum 'Bose' gases if the short-range interaction between their particles is attractive<sup>7</sup>. However, these solitons are unstable in 3D without confinement, whereas the authors show that their droplets are fully stable even without a trap. This advantage over solitons could mean that the droplets find practical applications, for example in atom interferometry, a measurement technique (based on the interference between quantum-mechanical waves of matter) in which self-bound objects could be used to increase the brightness of the interferometer<sup>8</sup>.

As long as confinement is maintained, Schmitt and collaborators' liquid droplets exist in equilibrium with atoms in a gaseous phase. However, one of the authors' most striking observations is that most of the energetic excitations rapidly leave the droplet after the trap has been removed — the two phases separate because, whereas the gas expands, the liquid remains localized. Although it is not known how cold



**Figure 1 | Quantum droplets.** Schmitt *et al.*<sup>2</sup> have created droplets of an ultracold gas of dysprosium atoms. **a**, The authors' research group showed<sup>3</sup> previously that such atoms, when contained in a trap, form an ordered arrangement of droplets if the dipole-dipole interactions between atoms are large enough relative to their short-range interactions. **b**, In the present paper, the authors produce single, self-bound droplets. In the trap, the liquid droplets coexist with atoms in the gaseous phase, but after the trap has been removed, the gas rapidly expands.

the authors' droplets are, this evaporation of the gas will probably result in the liquid having an extremely low temperature<sup>9</sup>. One could therefore envision applications in which the liquid state is used as a coolant for other quantum gases. For such applications, and others, it will be necessary to find a way to increase the lifetime of the droplets. For the moment, the lifetime is limited to about 0.1 seconds because of the spontaneous formation of molecules as a result of three-atom collisions. ■

**Bruno Laburthe-Tolra** is at the *Laboratoire de Physique des Lasers, Université Paris 13,*

*93430 Villetaneuse, France.*

*e-mail: laburthe-tolra@univ-paris13.fr*

1. Pitaevskii, L. & Stringari, S. *Bose-Einstein Condensation* (Clarendon, 2003).
2. Schmitt, M., Wenzel, M., Böttcher, F., Ferrier-Barbut, I. & Pfau, T. *Nature* **539**, 259–262 (2016).
3. Kadau, H. *et al. Nature* **530**, 194–197 (2016).
4. Dietrich, F., Chen, E., Quint, J. W. & Walter, H. *Phys. Rev. Lett.* **59**, 2931–2934 (1987).
5. Ferrier-Barbut, I. *et al. Phys. Rev. Lett.* **116**, 215301 (2016).
6. Chomaz, L. *et al.* Preprint at <https://arxiv.org/abs/1607.06613> (2016).
7. Strecker, K. E., Partridge, G. B., Truscott, A. G. & Hulet, R. G. *Nature* **417**, 150–153 (2002).
8. McDonald, G. D. *et al. Phys. Rev. Lett.* **113**, 013002 (2014).
9. Petrov, D. S. *Phys. Rev. Lett.* **115**, 155302 (2015).

#### SPINAL-CORD INJURY

## Neural interfaces take another step forward

**Two monkeys subjected to a spinal-cord injury that paralysed one leg have regained the ability to walk, thanks to technology that re-establishes communication between the brain and spinal cord. SEE LETTER P.284**

**ANDREW JACKSON**

Each time we decide to move, electrical signals cascade from the brain down the spinal cord to instruct our muscles to contract. Damage to the pathways that relay these messages, for example owing to a spinal-cord injury, can result in paralysis, for which there is currently no cure. The nerves of the spinal cord do not heal spontaneously after injury, and efforts to repair them with pharmacological and regenerative techniques have had only limited success. On page 284, Capogrosso *et al.*<sup>1</sup> report that an alternative approach involving a wireless electronic

connection between the brain and spinal cord restored movement in two monkeys that were each paralysed in one leg as a result of a spinal-cord injury.

The researchers used an implant in the brain called a neural interface, which decodes information from an array of electrodes that measure the activity of multiple brain cells that normally control leg movements (Fig. 1). They surgically implanted a second electrode array over the lumbar region in the lower part of each monkey's spinal cord, below the level of the injury — a partial lesion of the spinal cord that severed the nerve connections on one side of the cord. These electrodes delivered



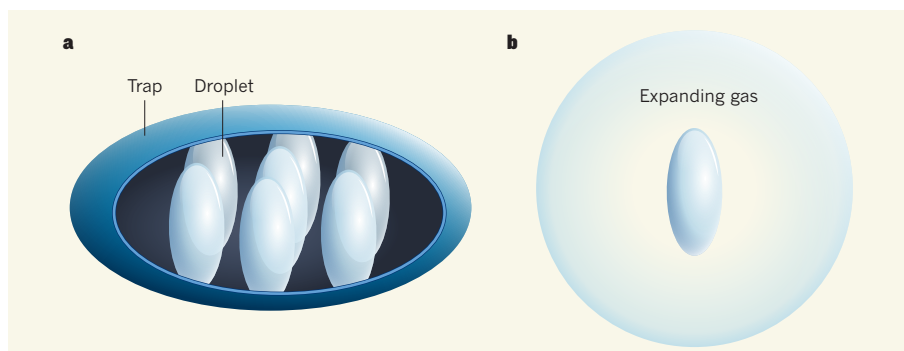
set-up. First, they produce a trap that is elongated in a direction parallel to the magnetic dipoles of the atoms. Instead of an ensemble of droplets, the atomic gas forms single, 'self-bound' droplets (Fig. 1b).

Schmitt and collaborators' second and decisive change is to remove the trap after the droplets have been produced, and to replace it with a simple magnetic-field gradient whose only effect is to compensate for gravity. This can be seen as an inexpensive way to perform the experiment in microgravity. Instead of expanding in space like a normal gas, the droplets remain localized. This is because the dipole-dipole interactions between the atoms provide attraction and cohesion, in a similar way to the hydrogen bonds between H<sub>2</sub>O molecules in water.

Perhaps one of the most interesting aspects of Schmitt and colleagues' work is their demonstration that there are a critical number of atoms below which the liquid droplets are unstable and turn into a gas. This phase transition is one of the first to be observed in a cold-atom experiment without a trap. To understand this transition, it is necessary to explain why the liquid droplets can be stable in the first place. The authors observe self-bound droplets when the balance between repulsive short-range interactions and attractive long-range interactions would have been expected to lead to collapse. However, collapse does not occur because van der Waals and dipole-dipole interactions create quantum fluctuations<sup>5,6</sup>. These fluctuations apply a pressure to the droplets that rapidly increases with density, and is therefore sufficient to ensure stability only if the number of atoms is large enough.

The main novelty of the authors' work is that their extremely dilute droplets are self-bound — a property that is reminiscent of solitonic structures. Solitons were named after non-dispersive, 'solitary' waves of water, but are well known in optics and also exist in the context of quantum 'Bose' gases if the short-range interaction between their particles is attractive<sup>7</sup>. However, these solitons are unstable in 3D without confinement, whereas the authors show that their droplets are fully stable even without a trap. This advantage over solitons could mean that the droplets find practical applications, for example in atom interferometry, a measurement technique (based on the interference between quantum-mechanical waves of matter) in which self-bound objects could be used to increase the brightness of the interferometer<sup>8</sup>.

As long as confinement is maintained, Schmitt and collaborators' liquid droplets exist in equilibrium with atoms in a gaseous phase. However, one of the authors' most striking observations is that most of the energetic excitations rapidly leave the droplet after the trap has been removed — the two phases separate because, whereas the gas expands, the liquid remains localized. Although it is not known how cold



**Figure 1 | Quantum droplets.** Schmitt *et al.*<sup>2</sup> have created droplets of an ultracold gas of dysprosium atoms. **a**, The authors' research group showed<sup>3</sup> previously that such atoms, when contained in a trap, form an ordered arrangement of droplets if the dipole-dipole interactions between atoms are large enough relative to their short-range interactions. **b**, In the present paper, the authors produce single, self-bound droplets. In the trap, the liquid droplets coexist with atoms in the gaseous phase, but after the trap has been removed, the gas rapidly expands.

the authors' droplets are, this evaporation of the gas will probably result in the liquid having an extremely low temperature<sup>9</sup>. One could therefore envision applications in which the liquid state is used as a coolant for other quantum gases. For such applications, and others, it will be necessary to find a way to increase the lifetime of the droplets. For the moment, the lifetime is limited to about 0.1 seconds because of the spontaneous formation of molecules as a result of three-atom collisions. ■

**Bruno Laburthe-Tolra** is at the *Laboratoire de Physique des Lasers, Université Paris 13,*

*93430 Villetaneuse, France.*

*e-mail: laburthe-tolra@univ-paris13.fr*

1. Pitaevskii, L. & Stringari, S. *Bose-Einstein Condensation* (Clarendon, 2003).
2. Schmitt, M., Wenzel, M., Böttcher, F., Ferrier-Barbut, I. & Pfau, T. *Nature* **539**, 259–262 (2016).
3. Kadau, H. *et al. Nature* **530**, 194–197 (2016).
4. Dietrich, F., Chen, E., Quint, J. W. & Walter, H. *Phys. Rev. Lett.* **59**, 2931–2934 (1987).
5. Ferrier-Barbut, I. *et al. Phys. Rev. Lett.* **116**, 215301 (2016).
6. Chomaz, L. *et al.* Preprint at <https://arxiv.org/abs/1607.06613> (2016).
7. Strecker, K. E., Partridge, G. B., Truscott, A. G. & Hulet, R. G. *Nature* **417**, 150–153 (2002).
8. McDonald, G. D. *et al. Phys. Rev. Lett.* **113**, 013002 (2014).
9. Petrov, D. S. *Phys. Rev. Lett.* **115**, 155302 (2015).

#### SPINAL-CORD INJURY

## Neural interfaces take another step forward

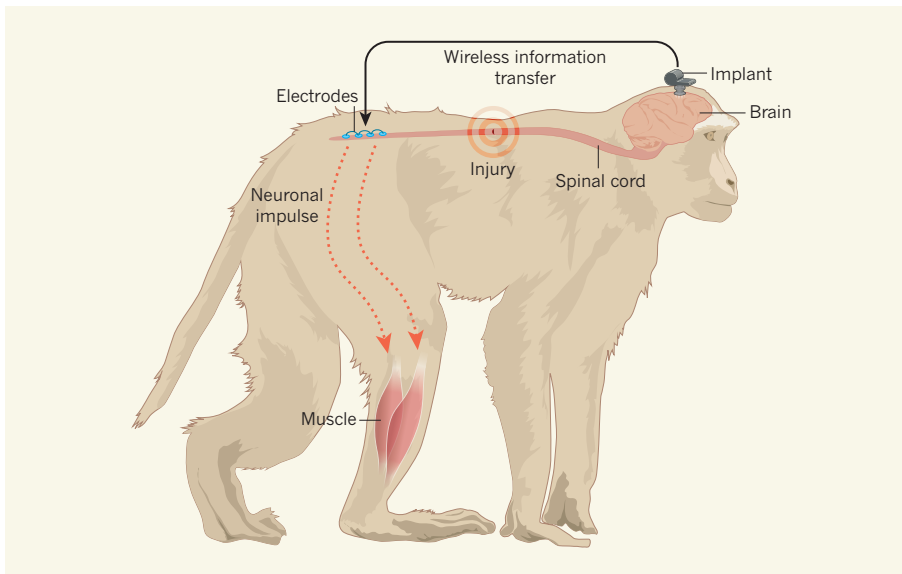
**Two monkeys subjected to a spinal-cord injury that paralysed one leg have regained the ability to walk, thanks to technology that re-establishes communication between the brain and spinal cord. SEE LETTER P.284**

**ANDREW JACKSON**

Each time we decide to move, electrical signals cascade from the brain down the spinal cord to instruct our muscles to contract. Damage to the pathways that relay these messages, for example owing to a spinal-cord injury, can result in paralysis, for which there is currently no cure. The nerves of the spinal cord do not heal spontaneously after injury, and efforts to repair them with pharmacological and regenerative techniques have had only limited success. On page 284, Capogrosso *et al.*<sup>1</sup> report that an alternative approach involving a wireless electronic

connection between the brain and spinal cord restored movement in two monkeys that were each paralysed in one leg as a result of a spinal-cord injury.

The researchers used an implant in the brain called a neural interface, which decodes information from an array of electrodes that measure the activity of multiple brain cells that normally control leg movements (Fig. 1). They surgically implanted a second electrode array over the lumbar region in the lower part of each monkey's spinal cord, below the level of the injury — a partial lesion of the spinal cord that severed the nerve connections on one side of the cord. These electrodes delivered



**Figure 1 | Communication bypass.** Capogrosso *et al.*<sup>1</sup> established a neural-interface system that enabled communication between the brain and spinal cord in monkeys subjected to a spinal-cord injury that paralysed one leg. An implant in the brain recorded neural activity related to leg movements, and transmitted this information to electrodes at the base of the spinal cord via a wireless link. These electrodes triggered neuronal impulses that generated movement in the leg muscle, allowing the monkeys to walk freely, despite their injury. (Adapted from Fig. 1 of ref. 1.)

small electrical currents that generated movements in the paralysed limb. Stimulation was controlled in real time by the decoded brain signals, which were relayed to the lumbar array through a wireless link, bypassing the injured region of the spine. This artificial connection restored communication between the brain and the lumbar spinal cord, enabling the animals to walk again.

The idea of using electronic implants to bypass damaged neural pathways dates back to the 1970s (ref. 2), but the twenty-first century has seen remarkable progress in this field. An early-stage clinical trial of a neural-interface technology called BrainGate showed<sup>3,4</sup> that brain signals recorded from humans who had been paralysed could facilitate control of a variety of devices, including computers and prosthetic limbs. Similar signals have also been used to control stimulation of the cervical region at the top of the spinal cord, to restore arm and hand movements in paralysed monkeys<sup>5</sup>. Moreover, lumbar spinal-cord stimulation has yielded promising results in human trials, in which participants with paraplegia have recovered some voluntary movement in their legs<sup>6</sup>. But at present, these human studies use only open-loop stimulation, which involves continuous trains of neural excitation — this strategy seems to boost the sensitivity of spinal locomotor circuits to weakened inputs from the brain that survived the injury.

The closed-loop stimulation used by Capogrosso *et al.*, in which neural stimulation is controlled in real time by brain signals, could enable more-accurate control of movements and perhaps work even in cases in which all neural inputs to the spinal cord have

been severed. Moreover, there is increasing evidence that closed-loop stimulation can drive neuroplasticity<sup>7</sup>, the mechanism by which the connections between two neurons strengthen if both are active at the same time. Neuroplasticity has a crucial role in rehabilitation following injuries that sever the spinal cord only partially. It is possible that, in such situations, stimulating the spinal cord when appropriate brain activity is detected could help to strengthen surviving motor pathways to promote lasting improvements in movement.

The pace at which neural interfaces are being translated from initial experiments in monkeys to human trials has been rapid. For example, just four years separated the first brain-controlled computer interfaces in monkeys<sup>8</sup> and people<sup>3</sup>. Similarly, a 2012 paper described the first use by a paralysed woman of a brain-controlled robotic arm<sup>4</sup> — a technology initially established in non-human primates only four years previously<sup>9</sup>. Earlier this year, the pattern was again repeated when voluntary hand-grasping was restored in a person with quadriplegia using brain-controlled muscle stimulation<sup>10</sup>, following monkey experiments<sup>11</sup> in 2012. It is therefore not unreasonable to speculate that we could see the first clinical demonstrations of interfaces between the brain and spinal cord by the end of the decade, especially because the implanted components used by Capogrosso *et al.* have already been approved for human use. This speed of translation from monkeys to humans is particularly impressive, given the technological and surgical complexity of neural interfaces, and it speaks to the close

anatomical and physiological similarities in the motor systems of all primates.

Despite this progress, the use of monkeys for neuroscience experiments continues to be questioned in the media, and animal-rights groups are making concerted efforts to ensure that restrictions on such work are tightened in both the United States and Europe. It is notable that, although Capogrosso and colleagues are based in Europe and their research conformed to the current regulations of the European Union, the experiments were conducted in China. Grégoire Courtine, the lead researcher on this study, has in the past described the challenges involved in performing such experiments abroad<sup>12</sup>, and other scientists might lack the time, energy or resources to pursue their research so far from home. There is thus a real danger that the development of treatments for debilitating neurological conditions will be delayed if high-quality, well-regulated research in monkeys cannot be performed in Europe and America owing to increasingly tight regulations. Equally, as more primate neuroscience moves to Asia, it will be important for researchers to remain committed to refining techniques and improving welfare standards for experimental animals worldwide.

The stakes are high. The World Health Organization estimates<sup>13</sup> that, each year, between 250,000 and 500,000 people suffer a spinal-cord injury, resulting in disabilities that can persist for decades. There are still key technological challenges to overcome before neural interfaces can record robust and stable brain signals over these long time periods. Moreover, it remains to be seen whether a brain–spinal-cord interface can restore bipedal walking in humans after injuries that affect both legs, the most common injury in the clinic. Useful locomotion also requires control of balance, steering and obstacle avoidance, which were not addressed in the current paper. Nevertheless, Capogrosso and colleagues' study represents a major step towards restoring lost motor function using neural interfaces. ■

**Andrew Jackson** is at the Institute of Neuroscience, Newcastle University, Newcastle NE2 4HH, UK.  
e-mail: andrew.jackson@ncl.ac.uk

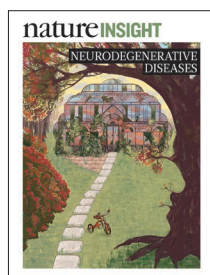
1. Capogrosso, M. *et al.* *Nature* **539**, 284–288 (2016).
2. Craggs, M. D. *Adv. Neurol.* **10**, 91–101 (1975).
3. Hochberg, L. R. *et al.* *Nature* **442**, 164–171 (2006).
4. Hochberg, L. R. *et al.* *Nature* **485**, 372–375 (2012).
5. Zimmermann, J. B. & Jackson, A. *Front. Neurosci.* **8**, 87 (2014).
6. Angeli, C. A., Edgerton, V. R., Gerasimenko, Y. P. & Harkema, S. J. *Brain* **137**, 1394–1409 (2014).
7. Jackson, A. & Zimmermann, J. B. *Nature Rev. Neurol.* **8**, 690–699 (2012).
8. Taylor, D. M., Helms Tillery, S. I. & Schwartz, A. B. *Science* **296**, 1829–1832 (2002).
9. Velliste, M., Perel, S., Spalding, M. C., Whitford, A. S. & Schwartz, A. B. *Nature* **453**, 1098–1101 (2008).
10. Bouton, C. E. *et al.* *Nature* **533**, 247–250 (2016).
11. Ethier, C., Oby, E. R., Bauman, M. J. & Miller, L. E. *Nature* **485**, 368–371 (2012).
12. Cyranoski, D. *Nature* **532**, 300–302 (2016).
13. [www.who.int/mediacentre/factsheets/fs384/en](http://www.who.int/mediacentre/factsheets/fs384/en)



# natureINSIGHT

## NEURODEGENERATIVE DISEASES

10 November 2016 / Vol 539 / Issue No 7628



Cover illustration  
Nik Spencer

**Editor, *Nature***  
Philip Campbell

**Publishing**  
Richard Hughes

**Insights Editor**  
Ursula Weiss

**Production Editor**  
Elizabeth Batty

**Art Editor**  
Nik Spencer

**Sponsorship**  
Janice Stevenson  
Samantha Meyer

**Production**  
Ian Pope

**Marketing**  
Nicole Jackson

**Editorial Assistant**  
Giacomo Russo

The Campus  
4 Crinan Street  
London N1 9XW, UK  
Tel: +44 (0) 20 7833 4000  
e: nature@nature.com

**SPRINGER  
NATURE**

The prevalence of neurodegenerative disorders is increasing, owing — in part — to extensions in lifespan. Currently, there is no cure for any of these diseases, although not for lack of trying. The hard work and dedication that goes into unravelling mechanisms of disease is discernible from this collection of reviews. Each summarises our knowledge, highlights exciting advances and provides ample inspiration for future research.

The signs of the passage of time are clearly visible in the brain. Tony Wyss-Coray synthesizes current knowledge on brain ageing and neurodegeneration and explores the prospect of stalling, or even resetting, the clock.

Growing evidence suggests that genetic, cellular and circuit dysregulation results from, and can lead to, cellular and cognitive hallmarks of Alzheimer's disease. Li-Huei Tsai, Rebecca Canter and Jay Penney argue for a multipronged approach to the treatment of this common form of dementia.

Paul Taylor, Robert Brown and Don Cleveland discuss emerging themes and mechanisms that underlie amyotrophic lateral sclerosis (also known as Lou Gehrig's disease or motor neuron disease), a progressive degeneration of motor neurons in the brain and spinal cord.

Parkinson's disease is characterized by the progressive death of dopamine neurons. Asa Abeliovich and Aaron Gitler propose that the accumulation of cellular damage eventually overwhelms the protein-disposal mechanisms of these neurons.

John Collinge considers the wider relevance of mammalian prions for neurodegenerative diseases. And Roland Riek and David Eisenberg provide a structural perspective on neurodegeneration through the properties of protein aggregates, the hallmarks of various neurodegenerative disorders. They explore the self-replication, cell-to-cell transmission and toxicity of these amyloids.

We hope that this collection will not only stimulate further research on neurodegenerative diseases but also direct more funding towards this area — as a greater understanding will reveal new opportunities for therapeutic intervention.

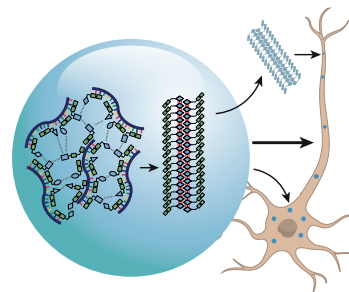
*Nature* is pleased to acknowledge the financial support of Eli Lilly and Company in producing this Insight. As always, *Nature* carries sole responsibility for all editorial content.

**Marie-Thérèse Heemels**  
*Senior Editor*

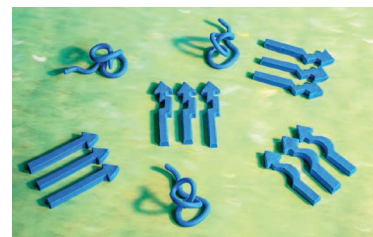
### CONTENTS

#### REVIEWS

- 180 Ageing, neurodegeneration and brain rejuvenation**  
Tony Wyss-Coray
- 187 The road to restoring neural circuits for the treatment of Alzheimer's disease**  
Rebecca G. Canter, Jay Penney & Li-Huei Tsai
- 197 Decoding ALS: from genes to mechanism**  
J. Paul Taylor, Robert H. Brown Jr & Don W. Cleveland



- 207 Defects in trafficking bridge Parkinson's disease pathology and genetics**  
Asa Abeliovich & Aaron D. Gitler
- 217 Mammalian prions and their wider relevance in neurodegenerative diseases**  
John Collinge
- 227 The activities of amyloids from a structural perspective**  
Roland Riek & David S. Eisenberg



# Ageing, neurodegeneration and brain rejuvenation

Tony Wyss-Coray<sup>1,2</sup>

**Although systemic diseases take the biggest toll on human health and well-being, increasingly, a failing brain is the arbiter of a death preceded by a gradual loss of the essence of being. Ageing, which is fundamental to neurodegeneration and dementia, affects every organ in the body and seems to be encoded partly in a blood-based signature. Indeed, factors in the circulation have been shown to modulate ageing and to rejuvenate numerous organs, including the brain. The discovery of such factors, the identification of their origins and a deeper understanding of their functions is ushering in a new era in ageing and dementia research.**

According to a 2015 United Nations report on world population ageing<sup>1</sup>, the number of people aged 60 and older worldwide is projected to more than double in the next 35 years, reaching almost 2.1 billion people. Most of this growth will come from developing regions of the world, although the oldest old, who are more than 80 years of age, are the fastest growing segment of the population in developed regions. Despite these improvements in life expectancy, Alzheimer's disease (AD) and related neurodegenerative conditions have arguably become the most dreaded maladies of older people. The observation that almost all aged brains show characteristic changes that are linked to neurodegeneration raises the question of whether these hallmarks represent lesser aspects of brain ageing that do not considerably affect function or whether they are the harbingers of neurodegenerative diseases (Fig. 1). Immune cells and secreted communication factors, which are responsible for tissue homeostasis in general, probably play important parts in brain ageing and neurodegeneration. However, comprehending or controlling the immune response in ageing has been a challenge. In the ageing organism, the brain seems to be susceptible to both cell-intrinsic and local signals, as well as to cues from the systemic environment. Animal models suggest that cues that are present in the circulatory system can either accelerate or slow aspects of brain ageing and cognitive function. This Review will synthesize present knowledge on brain ageing and neurodegeneration and discuss the prospect of stalling or even reversing these processes through circulatory factors.

## Overlap between ageing and neurodegeneration

Population-based autopsy studies of the brains of aged people who had not been diagnosed with a neurological disease consistently report the presence of amyloid plaques, neurofibrillary tangles, Lewy bodies, inclusions of TAR DNA-binding protein 43 (TDP-43), synaptic dystrophy, the loss of neurons and the loss of brain volume in most of the brains<sup>2</sup>. These features vary greatly between individuals, with particular lesions dominating a particular brain or restricted to specific regions. It is unknown what causes such lesions and whether they are the precursors to neurodegeneration and disease or simply the products of brain ageing. As well as classic protein deposits, other subcellular structures that consist of cross-linked proteins, carbohydrates or lipids accumulate in ageing brains, either in the extracellular space (for example, corpora amylacea) or inside glial cells or neuronal cells (for example, stress granules, lipofuscin, Marinesco bodies and Hirano bodies). Although

most of these structures are characterized poorly and their importance in neurodegeneration is unclear, it is probable that they take a toll on normal brain function<sup>3</sup>.

The presence of age-related protein abnormalities and inclusion bodies in the ageing brain points to defects in proteostasis, an idea that is supported by mounting evidence from experiments. According to one such hypothesis, in normal ageing, macromolecules become oxidized and can no longer be degraded by lysosomes<sup>4</sup>. This leads to the further production of lysosomal enzymes that are also unable to digest the cellular material. A well-known deposit that results from lysosomal inefficiency is lipofuscin, which is an accepted marker of ageing for postmitotic cells<sup>4</sup>. Similarly, the increase in damaged proteins and dying cells that accompanies ageing can overwhelm phagocytic processes and lead to an accumulation of material in lysosomes. Indeed, myelin debris have been demonstrated to accumulate in ageing microglia, in which it forms insoluble, lipofuscin-like lysosomal inclusions<sup>5</sup>. With ageing, and even more so with neurodegeneration, the brain shows increased levels of many lysosomal proteins and enzymes, and neurons and other cell types show abnormal endosomes, lysosomes and autophagosomes<sup>6–8</sup>. Whether these abnormalities contribute to or are the result of ageing must still be elucidated. However, the genetic manipulation of autophagy-related pathways in transgenic mice that overexpress amyloid precursor protein (also known as amyloid- $\beta$  A4 protein) or the protein tau results in prominent changes in the accumulation of these proteins or the progression of disease<sup>9–11</sup>. Furthermore, stress granules, which consist of RNA and protein and can form in response to cellular stress, might have an important role in amyotrophic lateral sclerosis and frontotemporal dementia<sup>12</sup>. Stress granules are also associated with aggregates of tau in the brains of people with AD and in the brains of mice that overexpress mutated tau protein, and the overexpression of stress-granule protein TIA1 seems to stimulate a tauopathy<sup>13</sup>. These studies underline the importance of protein homeostasis in brain function and suggest that ageing and neurodegeneration could result partly from a loss of proteostasis. Although cause and effect must again be established, it is probable that the stabilization of protein homeostasis would benefit the ageing brain.

Our limited knowledge about the relevance of protein abnormalities in the brain is demonstrated by a striking discrepancy between the clinical manifestations of dementia and its associated physical characteristics in the brain, particularly in the oldest old. For example, in a study

<sup>1</sup>Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Palo Alto, California 94304, USA. <sup>2</sup>Center for Tissue Regeneration, Repair and Restoration, VA Palo Alto Health Care System, Palo Alto, California 94304, USA.



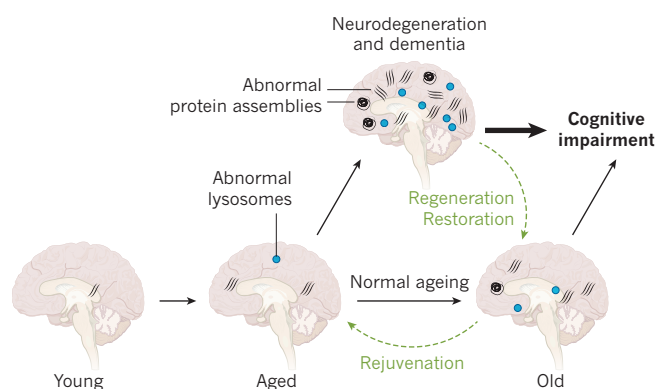
of a large series of brains from cognitively unimpaired aged people, almost all had abnormal accumulation of tau, roughly half had deposits of amyloid- $\beta$  or TDP43 and one-fifth had deposits of  $\alpha$ -synuclein<sup>2</sup>, although the regional distribution of these lesions should be considered when assessing their relevance to neurodegeneration. The brains of people who were aged 90 and older were found to weigh 11% less than those of individuals in their fifties<sup>2</sup>, which indicates that more than 150 g of brain tissue had vanished in the older brains. This disappearance could be due to the loss of neurons or glial cells, myelin, fluid or other factors, and it will be important to determine whether it is related to neurodegeneration or is simply a part of normal brain ageing. Similarly, in a population-based sample of nonagenarians and centenarians without dementia, almost half fulfilled the neuropathological criteria of AD or had a mix of numerous pathologies<sup>14</sup>. Yet of the nonagenarians and centenarians who had been clinically diagnosed with dementia, 12% were free of pathological features, 23% could be considered to have AD and 45% had mixed dementia<sup>14</sup>. These observations are supported by studies of cerebrospinal fluid biomarkers for amyloid- $\beta$  and tau, as well as positron-emission tomography imaging tracers in people, which show that around 30% of cognitively unimpaired elderly individuals are positive for these otherwise highly predictive markers of disease. Such individuals could be at a preclinical stage of AD, a stage of the disease that seems to be gaining clinical acceptance<sup>15</sup>. Around one quarter of cognitively healthy elderly people or people with mild cognitive impairment have pathological levels of tau in their brains in the absence of amyloid- $\beta$ , a condition called suspected non-AD pathophysiology (SNAP)<sup>16,17</sup>. Most such individuals do not express the apolipoprotein  $\epsilon 4$  (APOE4) isoform, which is consistent with the observation that APOE4 promotes the accumulation of amyloid- $\beta$  and that the APOE locus is linked to longevity.

In summary, clinical diagnoses often do not correlate with relevant pathological features in the brain, and there are few people above the age of 80 whose brains lack these features. The processes that characterize neurodegenerative diseases and, in particular, AD take place in most old brains; however, some people might have compensatory mechanisms that enable them to cope with these processes and to maintain normal cognition.

## Causes of brain ageing and neurodegeneration

Given that neurodegenerative diseases in the elderly are common and that disease-free brains, especially in the oldest old, are rare, it is possible that normal brain ageing forms a continuum with neurodegeneration and disease, and that stochastic factors, framed by a person's genetics and environment, determine the type of neurodegenerative disease that will dominate their brain eventually (Fig. 2). It is therefore tempting to view neurodegenerative diseases as expressions of accelerated ageing. However, this simplification is unhelpful because it does not accurately capture the underlying mechanisms that tie neurodegeneration to ageing, and all age-related diseases could essentially be described as forms of accelerated ageing. Instead, our understanding of how age contributes to disease is more likely to be advanced by dissecting how environmental factors and genes intersect in a particular disease with distinct hallmarks of ageing and by identifying the importance of these processes in the disease (Fig. 2). For example, lesions associated with a disease rather than ageing are often more region specific and cognitive changes with age seem to be distinct from those observed in AD<sup>18</sup>.

Twin studies show that the heritability of the human lifespan is 20–30% and that the genetic contribution increases with age<sup>19–21</sup>. The lower heritability at younger ages is probably caused by a greater number of accidental deaths at such ages<sup>22</sup>. Environmental factors therefore account for at least 70% of variation in lifespan and an increasing number of studies show that lifestyle, diet, exposure to toxins, including drugs of abuse, can have profound effects on healthspan, longevity and the development of neurodegenerative diseases, although the molecular pathways that underpin effects are mostly unknown. (The epidemiology of longevity is reviewed comprehensively elsewhere<sup>23</sup>.) Further insights

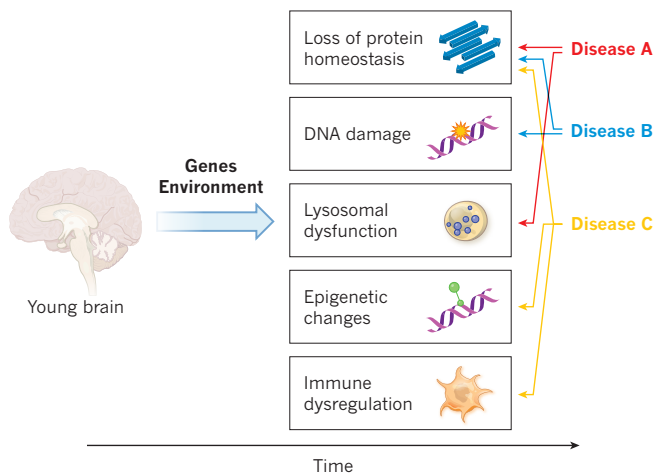


**Figure 1 | Ageing, neurodegeneration and brain rejuvenation.** As the brain ages, abnormal protein assemblies and inclusion bodies take hold and abnormal lysosomes are observed more frequently. It is unclear whether these defects promote ageing and neurodegeneration or whether they are innocent bystanders. Aged brains become highly prone to neurodegenerative diseases in which the same lesions amass as those that are found in old brains in smaller numbers. The relationship between such lesions and cognitive impairment is often blurred and normal aging and neurodegeneration and dementia can overlap. The concept of rejuvenation posits that old brains are malleable and that aspects of the ageing process can be reversed to a younger stage. If this can be achieved, it might also be possible to slow or reverse neurodegeneration and cognitive impairment.

into the links between ageing and neurodegeneration are being generated from genetic studies that explore not only longevity and exceptional lifespan but also the genetics of disease-free ageing<sup>24,25</sup> and the integration of genetics with other omics approaches<sup>26</sup>.

Exceptional longevity is linked consistently to the *TOMM40*–*APOE*–*APOC1* locus, and other strong links are observed at genes such as *FOXO3* and *IL6* (refs 21, 23 and 27). In a large meta analysis of centenarian cohorts, many of the single nucleotide polymorphisms (SNPs) linked to longevity with the greatest significance were linked negatively to AD and coronary heart disease<sup>28</sup>. Interestingly, healthy ageing – that is, ageing without developing a disease – does not seem to be linked to longevity genes; instead, it might be associated with the absence of risk factors for AD and cardiovascular disease<sup>24</sup>. In the same study, analysis of the genes of people aged 80 years or older who had not been affected by chronic diseases revealed links to SNPs that are involved in cognitive performance, which offers the possibility that brain health and cognition might be surrogates for or even determinants of healthy ageing.

Another approach to deciphering the mechanistic contribution of ageing to neurodegeneration examines the pace of ageing in a cell- and pathway-specific fashion that focuses on gene expression, DNA methylation and other epigenetic DNA modifications (Fig. 2) that change dramatically with age. For example, a comparison of age-related changes in gene expression in the brains of people with AD and those without the disease revealed that AD is characterized by signatures of accelerated ageing in a neuronal-stress gene expression module, which includes genes that are involved in protein folding and metabolism, and in an inflammation module, which is defined by genes involved in cytokines and microglia<sup>29</sup>. A strong positive correlation between ageing in various regions of the brain and methylation was observed in several hundred human brains<sup>30</sup>, a finding consistent with the epigenetic clock — a generalized DNA methylation pattern that seems to characterize most tissues<sup>31</sup>. Analysis of this pattern in Parkinson's disease showed that DNA methylation in blood cells is consistent with accelerated ageing<sup>32</sup>. A correlation of the pathology of AD with DNA methylation across the genome in almost 1,000 autopsied brains in two independent studies identified methylation sites close to the genes *ANK1*, *CDH23*, *RHBDF2* and *RPL13* that were linked to the disease<sup>33–35</sup>. Remarkably, all of these genes except for *CDH23* have biological links to the AD-associated gene *PTK2B*. A combination of transcriptome and epigenome analyses in



**Figure 2 | Cell-specific and pathway-specific acceleration of ageing.** Ageing can be dissected into individual processes, including a loss of protein homeostasis that leads to the development of aggregates and inclusion bodies, DNA damage, lysosomal dysfunction, epigenetic changes and immune dysregulation. The genetic predisposition of an individual, together with his or her exposure to the environment, determine the incidence and prevalence of the lesions that result from such processes, probably in a cell-specific manner. Various diseases might develop in accordance with the spatiotemporal distribution of the lesions.

brains affected by AD and in a mouse model of AD-related neurodegeneration enabled the discovery of a downregulation of genes and regulatory regions involved in synaptic plasticity as well as a concomitant increase in the expression of genes involved in immune response and regulatory regions<sup>36</sup> — most notably *SPI1*, which encodes PU.1, a transcription factor with importance for the development of the myeloid lineage, including microglia<sup>37</sup>.

Genetic and epigenetic studies can therefore help to uncover the molecular pathways that link ageing with neurodegeneration. More refined omics studies, conducted with single cells isolated from defined brain regions, will probably deepen our insights and enable us to identify new targets to delay aspects of ageing in a disease-specific fashion.

### The circulatory proteome of organismal ageing

Through plastic surgery, people can look years younger than their wrinkled hands, and although stretching the skin might not change its intrinsic age, it poses the question of whether all of a person's organs age at a similar pace. Ageing has been categorized into nine separate processes or hallmarks<sup>38</sup>, seven of which consist of molecular, mostly cell-intrinsic, changes such as telomere shortening, mitochondrial dysfunction or DNA damage. The other 'integrative' hallmarks include stem-cell ageing and dysfunction of intercellular communication<sup>38</sup>. The molecular or pharmacological manipulation of several of these processes has been shown to affect lifespan in mammals.

From an organismal perspective, intercellular communication is of particular interest as it could provide insights into the ageing process as well as help to identify biomarkers of ageing. Cellular communication occurs at numerous levels, from cells to tissues, across an organism and is accomplished by a myriad of molecules, including secreted proteins, lipids and metabolites, which must be tightly controlled. This network of communication factors changes as an organism develops, ages or is affected by disease. It is possible that age-related changes in cellular communication are simply adaptations to ageing. However, such changes might also contribute to ageing, either locally or distantly, and as a consequence, a particular organ or cell type might modulate or even control ageing at the organismal level (Fig. 3).

Technologies for studying the proteome, the lipidome and the metabolome can be used to characterize age-related changes, and an increasing number of studies are describing changes in the blood that

occur with normal brain ageing or with neurodegeneration, under the hypothesis that such changes mirror, in part, changes in the brain. As there are few studies of blood-based lipids or metabolites that regulate brain function, I will focus instead on studies that quantify secreted signalling proteins involved in intercellular communication — a subset of the proteome that has been dubbed the 'communicome'<sup>39</sup>. The most comprehensive study of the cellular communicome of ageing quantified around 1,100 proteins using aptamer-based assays of the blood of about 800 people<sup>40</sup>. The protein most strongly correlated with ageing, chordin-like protein 1, is an antagonist of bone morphogenetic protein 4 and might therefore be involved in neural stem-cell fate and angiogenesis. Other proteins with links to ageing include pleiotrophin, which is a neurotrophic and mitogenic factor, the metalloproteinase inhibitor TIMP1 and the cysteine-proteinase inhibitor cystatin-C, all of which were also associated strongly with ageing in human cerebrospinal fluid<sup>41</sup>. Of the 281 proteins detected in cerebrospinal fluid, 81 were correlated significantly with age in 90 cognitively unimpaired people aged between 21 and 85 (ref. 41), which indicates that the brain is exposed to very different environments depending on the expression of these communication factors. Using antibody-based multiplex assays, several studies have measured tens to hundreds of known communication factors in the blood plasma of people with various stages of AD, reporting protein signatures that characterize the prodromal stages of the disease<sup>42</sup> or the progression from early to late-stage AD<sup>39,43,44</sup>. Other studies have described protein signatures that correlate with *APOE* genotypes<sup>45</sup> or with cerebrospinal-fluid levels of amyloid- $\beta$  and tau in people with AD<sup>46,47</sup>. The aptamer platform, which is more encompassing and precise than alternative methods, was also used to measure 1,001 proteins in almost 700 people with no cognitive impairment, mild cognitive impairment or AD: 14% of proteins showed a significant association with AD and 13 proteins could be used to classify AD with 70% accuracy<sup>48</sup>. By combining plasma proteomic data from healthy individuals, AD and frontotemporal dementia with existing brain gene-expression data and data from genome-wide association studies (GWAS), the most prominent changes in AD were found to relate to TGF- $\beta$ -BMP-GDF signalling, the activation of complement and apoptosis, and GDF-3 was linked to neurogenesis and AD<sup>49</sup>. Although common factors, including *APOE*, complement, CCL5, clusterin and ICAM1, have been identified in these studies, it will be crucial to replicate them independently and to establish the *in vivo* biological importance of newly discovered proteins. If validated, such proteins or their combinations could become useful markers for brain ageing or neurodegeneration, as well as potential therapeutic targets.

To address this challenge, several communication factors were measured in the plasma of both young and aged mice, as well as in mice exposed to the blood of young or aged mice through heterochronic parabiosis, using an antibody-based multiplex assay<sup>50</sup>. (Parabiosis is established by surgically joining two mice at their flanks, leading to the formation of a vascular anastomosis and a shared circulatory system.) One of the factors that correlated most strongly with ageing and the effects of parabiosis on hippocampal neurogenesis was eotaxin (also known as CCL11), a small chemokine with a role in allergies and certain types of parasitic infections. Indeed, systemic administration of recombinant CCL11 to young mice was sufficient to reduce neurogenesis and to impair cognition (Table 1). In line with these potentially detrimental effects on the brain and cognition, the level of CCL11 increases in the choroid plexus during ageing<sup>51</sup> and in fat deposits with obesity, and it decreases after exercise in people who are obese<sup>52,53</sup>. Similarly,  $\beta$ -2-microglobulin (B2M), a component of major histocompatibility complex class I (MHC I) molecules, was found to be a pro-ageing factor that can impair cognition and neurogenesis in young mice and is necessary to maintain these functions in old mice<sup>54</sup> (Table 1). Together with studies that link MHC I molecules with synaptic plasticity and brain repair<sup>55</sup>, these findings further implicate the MHC I locus in brain ageing and neurodegeneration. Importantly, these studies also indicate that proteins in the blood circulation



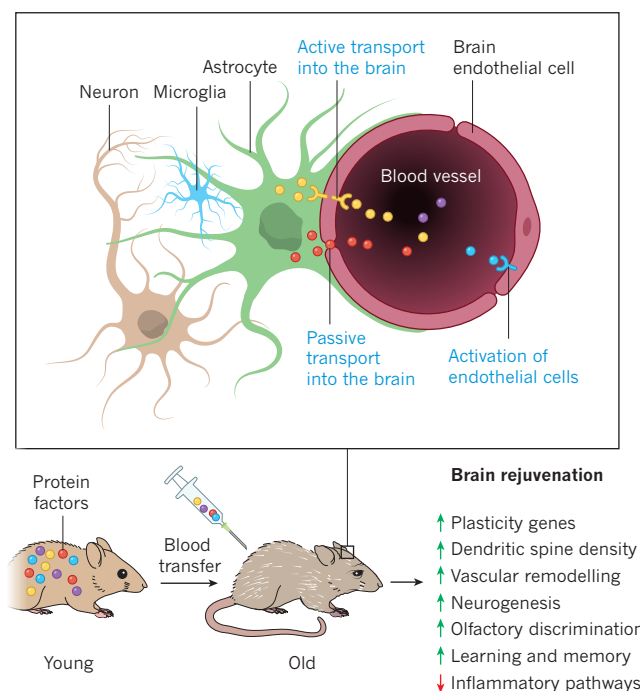
involved in intercellular communication are both correlated with and able to modulate brain ageing, and they demonstrate the feasibility of using plasma proteomics to discover factors of relevance to brain ageing and cognitive function.

### Abnormal intercellular communication and inflammation

Interestingly, the immune factors CCL11 and B2M, as well as the chemokines CCL2, CCL12 and CCL19 and haptoglobin, which are linked to negative effects on neurogenesis during parabiosis<sup>50</sup>, might be part of a low-grade inflammation that is linked to ageing, known as inflammageing<sup>56</sup>. Inflammatory factors in the ageing brain could originate from microglia and astrocytes as they become senescent and adopt a senescence-associated secretory phenotype<sup>57</sup>. Some ageing astrocytes express increased levels of cytokines, intermediate filament proteins and intracellular protein aggregates, which is consistent with the phenotype<sup>58</sup>. As discussed previously, this senescent phenotype could result from epigenetic changes that activate immune-response genes targeted by, for example, the transcription factor PU.1 (ref. 36). Alternatively, microglia that change their gene-expression repertoire dramatically with age in a brain-region-specific manner<sup>59</sup> might become reactive and inflamed as a result of impaired phagocytosis and protein dyshomeostasis. (A detailed discussion of microglia in brain ageing is presented elsewhere<sup>60</sup>.)

A role for inflammatory factors in autosomal dominant forms of neurodegeneration was suggested by the observation that numerous SNPs in the chemokine cluster that contains the gene *CCL11* were linked to a 10-year difference in the age of onset of clinical AD symptoms in 72 people carrying a highly penetrant presenilin 1 mutation<sup>61</sup>. Indeed, inflammation has long been associated with neurodegeneration<sup>62,63</sup>, and the use of non-steroidal anti-inflammatory drugs for several years before the onset of clinical symptoms is associated with a reduced risk of AD<sup>64–66</sup>. However, the same drugs do not seem to benefit people with Parkinson's disease<sup>67</sup>. It is unclear exactly how inflammation contributes to AD but it might involve both local and systemic mechanisms. In support of a detrimental role for systemic inflammation in the early stages of AD, the number of systemic inflammatory events (such as urinary tract infections) correlate positively with the progression and severity of AD<sup>68,69</sup>. Genome-wide transcriptome studies of numerous brain regions in more than 1,600 brains provide further evidence of a role for immune mechanisms in ageing and AD; they also show that the expression of genes involved in inflammation increases considerably with normal ageing and precedes the development of AD<sup>29</sup>. Another bioinformatics-based study used signalling pathway and network analysis to conclude that the gene *TYROBP* (also known as *DAP12*), restricted mainly to microglia in the brain, is deregulated in AD<sup>70</sup>. *DAP12* is an adaptor for several receptor molecules, including complement receptor 3, an important phagocytic receptor expressed by microglia, and *TREM2* (ref. 71). The most direct evidence that altered immune function has a role in AD emerged from genetic studies that showed that rare polymorphisms in the myeloid-lineage gene *TREM2* increase the risk of developing AD several fold<sup>72,73</sup>. GWAS also identified further polymorphisms in genes involved in immune responses that modify the risk of developing AD<sup>74–76</sup>. In the brain, most of these genes, including *TREM2*, are expressed predominantly or exclusively by microglia. Dysfunction of microglia would probably impair the capacity of these cells to uptake and degrade amyloid- $\beta$  and could therefore directly promote or even initiate AD. Antibodies that bind amyloid- $\beta$  to facilitate its clearance by microglia are being tested in the clinic at present<sup>77</sup>.

Together, the genetic, transcriptomic and proteomic evidence suggests that changes in inflammation and intercellular communication represent chief aspects of normal brain ageing and neurodegeneration. However, it is unclear whether inflammatory pathways simply drive ageing and disease or whether aspects of the inflammatory response fulfil reparative and regenerative functions.



**Figure 3 | Brain rejuvenation through circulatory factors from a young mouse.** Protein factors and other molecules that circulate in the blood of a young mouse exert rejuvenating effects on the brain of an old mouse after intravenous delivery. Such factors might affect the brain through various mechanisms, including: the active transport of factors into the brain parenchyma; the passive transport or diffusion of factors into the brain; and the activation of endothelial cells by factors. Other mechanisms might include the uptake of factors into the brain in areas that lack the blood–brain barrier, such as the neurogenic subventricular zone and circumventricular organs (not shown). As well as exerting effects on the endothelium, factors that enter the brain parenchyma might act on neurons, microglia or oligodendrocytes (not shown). The reported rejuvenating effects of young blood or young plasma on the old brain are listed.

### Brain rejuvenation and the manipulation of ageing

The concept of organismal and systemic ageing has been tested radically using heterochronic parabiosis<sup>78</sup>, which enables the exchange of blood, including its cells and factors, between young and old organisms. This surgically and conceptually simple model can therefore be used to investigate whether a youthful intercellular communicome can inhibit or reverse age-related abnormalities in an old mouse or whether an aged, and possibly dysfunctional, communicome can promote ageing in a young mouse. According to studies from nine independent laboratories, stem-cell activity is increased and other indices of ageing are delayed or reversed in several tissues of aged mice that share a circulatory system with young mice. These studies included: initial observations of effects on muscle and the liver<sup>78</sup>; reports of effects on the brain by four separate laboratories<sup>50,79–81</sup> (Table 1); and observations of rejuvenating effects in the pancreas<sup>82</sup>, the heart<sup>83</sup>, bone<sup>84</sup> and muscle<sup>85</sup> (for detailed reviews, see refs 86 and 87). By contrast, the ageing thymus does not seem to benefit from parabiosis; however, the injection of young epithelial cells enabled thymic regrowth<sup>88</sup>. Perhaps most remarkable, with respect to the brain, is that the repeated intravenous administration of plasma, the soluble fraction of blood, from young mice (performed systematically for the first time in 2011, to study ageing factors<sup>50</sup>), was sufficient to improve cognitive function in old mice in several behavioural tests<sup>80</sup> (Fig. 3). These functional changes were accompanied by molecular, subcellular, cellular and electrophysiological correlates, which suggests that factors in young blood have the capacity both to regulate brain function and to improve it to levels found in younger mice (Table 1). Parabiosis of amyloid-precursor protein transgenic mice with young mice, or the

**Table 1 | Effects of systemically administered ageing and rejuvenation factors on healthy brains**

Factors	Organism or model	Effect on the brain	References
Blood or plasma from old mouse	Young adult mouse	In the young brain: reduction in neurogenesis; increase in microglial reactivity; reduction in learning and memory.	50
CCL11	Young adult mouse	In the young brain: reduction in neurogenesis; increase in microglial reactivity; reduction in learning and memory.	50
B2M	Young adult mouse	In the young brain: reduction in neurogenesis; increase in microglial reactivity; reduction in learning and memory.	54
Blood or plasma from young mouse	Aged mouse	Increase in neurogenesis; reduction in microglial reactivity; improvement in learning and memory; improvement in olfactory discrimination.	80, 81
IGF1	Young adult rat	Increase in neurogenesis.	91
IGF2	Young adult mouse	Increase in retention and persistence of working, short-term and long-term memory.	92
GHRH	Healthy older people; people with mild cognitive impairment	A 20-week treatment improved executive functions in both groups of people.	94
GnRH	Aged mouse	Increase in neurogenesis; improvement in memory.	95
GDF-11	Aged mouse	Increase in neurogenesis; increase in cerebrovascular integrity.	81

IGF, insulin-like growth factor; GHRH, growth hormone-releasing hormone; GnRH, gonadotropin-releasing hormone; GDF-11, growth and differentiation factor 11.

intravenous administration of plasma from young mice, also reversed the loss of synaptophysin and calbindin (an indicator of cognitive decline both in people with AD and in transgenic mouse models of the disease), normalized MAPK–ERK signalling and improved their working memory<sup>89</sup>.

A deficiency in growth hormone or insulin–insulin-like growth factor I signalling has been linked to deficits in memory, and the activation of growth hormone and insulin-like growth factor I signalling has been associated with improved brain function after injury or under conditions in which these proteins are lacking (reviewed in ref. 90). Few studies have treated healthy aged animals or people with growth factors and even fewer have demonstrated beneficial effects of such treatment (Table 1). For example, although insulin-like growth factor I increases neurogenesis<sup>91</sup> and insulin-like growth factor II increases memory<sup>92</sup> in young rodents, these factors have not been tested in aged animals. And in healthy older women or women with mild cognitive impairment, treatment with insulin-like growth factor I for 1 year showed no effect on bone density, bone strength, mood or memory<sup>93</sup>. By contrast, systemic administration of growth hormone-releasing hormone (GHRH), which triggers the hypothyseal release of growth hormone and increases the levels of circulating insulin-like growth factor I, among other factors, resulted in improved cognition in healthy elderly people or in people with mild cognitive impairment<sup>94</sup>. On the basis of the observation that the hypothalamus might have a role in regulating organismal ageing, systemic treatment of aged mice with gonadotropin-releasing hormone I (GnRH I) was found to increase neurogenesis and to improve cognitive function<sup>95</sup>.

In a search for factors that decrease with ageing and that might be responsible for the beneficial effects of heterochronic parabiosis on the heart, GDF-11 was identified as a potential heart-rejuvenation factor<sup>83</sup>. Subsequently, GDF-11 was found to increase neurogenesis, to improve olfaction and to exert beneficial effects on the brain vasculature<sup>81</sup> (Table 1), as well as on aged muscle<sup>85</sup>. Other studies were unable to repeat these effects on systemic tissues<sup>96</sup>, and mass-spectrometry-based assays for GDF-11 and the related protein myostatin (GDF-8) observed no decrease in GDF-11 levels in human plasma with ageing and also found that GDF-11 levels are associated with frailty in people with cardiovascular disease<sup>97</sup>. Further studies will need to determine whether particular forms of GDF-11 (for example, mature, immature or post-translationally modified GDF-11) can explain these discrepancies and, most importantly, whether systemic administration of GDF-11 might be beneficial for human brains.

Overall, parabiosis with young mice or the transfer of young plasma seems to be capable of restoring brain function in old mice to more

youthful levels. GHRH, GnRH I and GDF-11 are putative brain-rejuvenation factors and it is probable that other age-related proteins with detrimental or beneficial effects on the brain will be discovered. So far, it is unknown how the plasma from young animals or the factors listed in Table 1 exert their effects. It is possible that some of these proteins enter the brain actively or passively through the blood–brain barrier or at sites that lack a functional barrier, including the circumventricular organs and, perhaps, the neurogenic niches (Fig. 3). Other proteins might modulate vascular function by interacting with endothelial cells and modulating the neurovascular unit<sup>81</sup>. In the future, studies will have to determine these modes of action and explore their potential for use as therapeutic approaches.

## Outlook

In humans, the old brain shows the classic hallmarks of ageing and is particularly susceptible to abnormal protein accumulation and impairments in the phagolysosomal system, which leads to fluid boundaries between ageing and neurodegenerative diseases. Consequently, many old people have pathological abnormalities of the brain that do not necessarily correlate with their cognitive abilities. This has important implications for the treatment of those with clinical symptoms as well as for designing clinical trials to target protein abnormalities in a specific manner. Given the crucial functions that immune responses and inflammation have in brain ageing and neurodegeneration, it will be essential to discern beneficial attempts to maintain or repair damage from maladaptive ones. Clearly, the term neuroinflammation fails to capture the age- or disease-related changes in this sophisticated interplay between the surveillance, identification, targeting and execution functions of immunity and should probably be avoided. When studying age-related neurodegenerative diseases in animal models, it is important to consider ageing; those that have been genetically engineered to develop disease during adolescence and before midlife are unlikely to be influenced sufficiently by ageing and are therefore not very informative about age-related factors in sporadic neurodegeneration.

The increasing number of studies that show systemic effects on the brain, including those of young plasma or heterochronic parabiosis, as well as the effects of the microbiome, should remind neuroscientists that neurons do not function in isolation; instead, they are part of a sophisticated network that includes glial cells, vascular cells and peripheral cells.

So far, there is no published evidence that young blood or plasma has beneficial effects on an ageing human body, and the observation that young plasma can modulate brain ageing in mice presents more questions and opportunities than answers. Only a handful of proteins, which might represent factors involved in ageing or rejuvenation, have been



shown to mimic the effects of plasma. However, many more proteins or other types of molecules are likely to exist, some of which might have direct therapeutic applications. Basic research will address the exciting questions that surround the origins of these factors, how they signal to the brain and why they change with age. Ultimately, it is hoped that by using such knowledge to alter basic processes involved in ageing, it will become feasible to counter the cellular abnormalities that lead to neurodegeneration. ■

Received 3 May; accepted 2 September 2016.

1. United Nations Department of Economic and Social Affairs, Population Division. *World Population Ageing 2015*. Report ST/ESA/SER.A/390 [http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015\\_Report.pdf](http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf) (United Nations, 2015).
2. Eloheid, A., Libard, S., Leino, M., Popova, S. N. & Alafuzoff, I. Altered proteins in the aging brain. *J. Neuropathol. Exp. Neurol.* **75**, 316–325 (2016).  
**A comprehensive study and review of the literature describing the prevalence of protein aggregates in cognitively unimpaired aged brains.**
3. Mrazek, R. E., Griffin, S. T. & Graham, D. I. Aging-associated changes in human brain. *J. Neuropathol. Exp. Neurol.* **56**, 1269–1275 (1997).
4. Brunk, U. T. & Terman, A. The mitochondrial-lysosomal axis theory of aging: accumulation of damaged mitochondria as a result of imperfect autophagocytosis. *Eur. J. Biochem.* **269**, 1996–2002 (2002).
5. Safaiyan, S. *et al.* Age-related myelin degradation burdens the clearance function of microglia during aging. *Nature Neurosci.* **19**, 995–998 (2016).  
**An elegant demonstration of how age-related myelin breakdown results in the accumulation of microglial lipofuscin and cell dysfunction.**
6. Nixon, R. A., Cataldo, A. M. & Mathews, P. M. The endosomal-lysosomal system of neurons in Alzheimer's disease pathogenesis: a review. *Neurochem. Res.* **25**, 1161–1172 (2000).
7. Nixon, R. A. *et al.* Extensive involvement of autophagy in Alzheimer disease: an immuno-electron microscopy study. *J. Neuropathol. Exp. Neurol.* **64**, 113–122 (2005).
8. Menzies, F. M., Fleming, A. & Rubinsztein, D. C. Compromised autophagy and neurodegenerative diseases. *Nature Rev. Neurosci.* **16**, 345–357 (2015).
9. Pickford, F. *et al.* The autophagy-related protein beclin 1 shows reduced expression in early Alzheimer disease and regulates amyloid  $\beta$  accumulation in mice. *J. Clin. Invest.* **118**, 2190–2199 (2008).
10. Nilsson, P. *et al.*  $A\beta$  secretion and plaque formation depend on autophagy. *Cell Rep.* **5**, 61–69 (2013).
11. Caccamo, A. *et al.* mTOR regulates tau phosphorylation and degradation: implications for Alzheimer's disease and other tauopathies. *Aging Cell* **12**, 370–380 (2013).
12. Ash, P. E. A., Vanderweyde, T. E., Youmans, K. L., Picco, D. J. & Wolozin, B. Pathological stress granules in Alzheimer's disease. *Brain Res.* **1584**, 52–58 (2014).
13. Vanderweyde, T. *et al.* Contrasting pathology of the stress granule proteins TIA-1 and G3BP in tauopathies. *J. Neurosci.* **32**, 8270–8283 (2012).
14. Kwas, C. H. *et al.* Multiple pathologies are common and related to dementia in the oldest-old: The 90+ Study. *Neurology* **85**, 535–542 (2015).
15. Dubois, B. *et al.* Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. *Alzheimers Dement.* **12**, 292–323 (2016).
16. Jack, C. R. *et al.* Suspected non-Alzheimer disease pathophysiology — concept and controversy. *Nature Rev. Neurol.* **12**, 117–124 (2016).
17. Landau, S. M., Horng, A., Fero, A. & Jagust, W. J. Amyloid negativity in patients with clinically diagnosed Alzheimer disease and MCI. *Neurology* **86**, 1377–1385 (2016).
18. Small, S. A., Schobel, S. A., Buxton, R. B., Witter, M. P. & Barnes, C. A. A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nature Rev. Neurosci.* **12**, 585–601 (2011).
19. Herskind, A. M. *et al.* The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900. *Hum. Genet.* **97**, 319–323 (1996).
20. vB Hjelmborg, J. *et al.* Genetic influence on human lifespan and longevity. *Hum. Genet.* **119**, 312–321 (2006).
21. Shadyab, A. H. & LaCroix, A. Z. Genetic factors associated with longevity: a review of recent findings. *Ageing Res. Rev.* **19**, 1–7 (2015).
22. Pilling, L. C. *et al.* Human longevity is influenced by many genetic variants: evidence from 75,000 UK Biobank participants. *Aging* **8**, 547–560 (2016).
23. Newman, A. B. & Murabito, J. M. The epidemiology of longevity and exceptional survival. *Epidemiol. Rev.* **35**, 181–197 (2013).
24. Erikson, G. A. *et al.* Whole-genome sequencing of a healthy aging cohort. *Cell* **165**, 1002–1011 (2016).
25. Matteini, A. M. *et al.* GWAS analysis of handgrip and lower body strength in older adults in the CHARGE consortium. *Aging Cell* **15**, 792–800 (2016).
26. Putin, E. *et al.* Deep biomarkers of human aging: application of deep neural networks to biomarker development. *Aging* **8**, 1021–1033 (2016).
27. Zeng, Y. *et al.* Novel loci and pathways significantly associated with longevity. *Sci. Rep.* **6**, 21243 (2016).
28. Sebastiani, P. *et al.* Meta-analysis of genetic variants associated with human exceptional longevity. *Aging* **5**, 653–661 (2013).
29. Podtelezniukov, A. A. *et al.* Molecular insights into the pathogenesis of Alzheimer's disease and its relationship to normal aging. *PLoS ONE* **6**, e29610 (2011).
30. Hernandez, D. G. *et al.* Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Hum. Mol. Genet.* **20**, 1164–1172 (2011).
31. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013); erratum **16**, 96 (2015).  
**Description of an 'epigenetic clock' that correlates with tissue ageing and shows acceleration in cancer.**
32. Horvath, S. & Ritz, B. R. Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. *Aging* **7**, 1130–1142 (2015).
33. Lunnon, K. *et al.* Methylomic profiling implicates cortical deregulation of *ANK1* in Alzheimer's disease. *Nature Neurosci.* **17**, 1164–1170 (2014).
34. De Jager, P. L. *et al.* Alzheimer's disease: early alterations in brain DNA methylation at *ANK1*, *BIN1*, *RHBDF2* and other loci. *Nature Neurosci.* **17**, 1156–1163 (2014).  
**A large-scale genome-wide DNA methylation study of a neurodegenerative disease.**
35. Lord, J. & Cruchaga, C. The epigenetic landscape of Alzheimer's disease. *Nature Neurosci.* **17**, 1138–1140 (2014).
36. Gjonneska, E. *et al.* Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* **518**, 365–369 (2015).
37. Satoh, J.-I., Asahina, N., Kitano, S. & Kino, Y. A comprehensive profile of ChIP-Seq-based PU.1/Spi1 target genes in microglia. *Gene Regul. Syst. Bio.* **8**, 127–139 (2014).
38. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194–1217 (2013).
39. Ray, S. *et al.* Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nature Med.* **13**, 1359–1362 (2007).
40. Menni, C. *et al.* Circulating proteomic signatures of chronological age. *J. Gerontol. A Biol.* **70**, 809–816 (2015).
41. Baird, G. S. *et al.* Age-dependent changes in the cerebrospinal fluid proteome by slow off-rate modified aptamer array. *Am. J. Pathol.* **180**, 446–456 (2012).
42. Hu, W. T. *et al.* Plasma multianalyte profiling in mild cognitive impairment and Alzheimer disease. *Neurology* **79**, 897–905 (2012); erratum **79**, 1935 (2012).
43. Johnstone, D., Milward, E. A., Berretta, R. & Moscat, P. Multivariate protein signatures of pre-clinical Alzheimer's disease in the Alzheimer's Disease Neuroimaging Initiative (ADNI) Plasma Proteome Dataset. *PLoS ONE* **7**, e34341 (2012).
44. Hye, A. *et al.* Plasma proteins predict conversion to dementia from prodromal disease. *Alzheimers Dement.* **10**, 799–807 (2014).
45. Soares, H. D. *et al.* Plasma biomarkers associated with the apolipoprotein E genotype and Alzheimer disease. *Arch. Neurol.* **69**, 1310–1317 (2012).
46. Britschgi, M. *et al.* Modeling of pathological traits in Alzheimer's disease based on systemic extracellular signaling proteome. *Mol. Cell. Proteomics* **10**, M111.008862 (2011).
47. Kiddle, S. J. *et al.* Plasma based markers of [ $^{11}\text{C}$ ] PiB-PET brain amyloid burden. *PLoS ONE* **7**, e44260 (2012).
48. Sattler, M. *et al.* Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. *Alzheimers Dement.* **10**, 724–734 (2014).
49. Jaeger, P. A. *et al.* Network-driven plasma proteomics expose molecular changes in the Alzheimer's brain. *Mol. Neurodegener.* **11**, 31 (2016); erratum **11**, 42 (2016).
50. Villeda, S. A. *et al.* The ageing systemic milieu negatively regulates neurogenesis and cognitive function. *Nature* **477**, 90–94 (2011).  
**The first demonstration of the effects of circulatory blood factors on brain ageing and cognitive function and the first systematic treatment of ageing with plasma injections.**
51. Baruch, K. *et al.* CNS-specific immunity at the choroid plexus shifts toward destructive Th2 inflammation in brain aging. *Proc. Natl Acad. Sci. USA* **110**, 2264–2269 (2013).
52. Vasudevan, A. R. Eotaxin and obesity. *J. Clin. Endocrinol. Metab.* **91**, 256–261 (2006).
53. Choi, K. M. *et al.* Effect of exercise training on plasma visfatin and eotaxin levels. *Eur. J. Endocrinol.* **157**, 437–442 (2007).
54. Smith, L. K. *et al.*  $\beta$ 2-microglobulin is a systemic pro-aging factor that impairs cognitive function and neurogenesis. *Nature Med.* **21**, 932–937 (2015).
55. Shatz, C. J. MHC class I: an unexpected role in neuronal plasticity. *Neuron* **64**, 40–45 (2009).
56. Franceschi, C. & Campisi, J. Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. *J. Gerontol. A Biol.* **69**, S4–S9 (2014).
57. Coppé, J.-P. *et al.* Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.* **6**, 2853–68 (2008).
58. Salminen, A. *et al.* Astrocytes in the aging brain express characteristics of senescence-associated secretory phenotype. *Eur. J. Neurosci.* **34**, 3–11 (2011).
59. Grabert, K. *et al.* Microglial brain region-dependent diversity and selective regional sensitivities to aging. *Nature Neurosci.* **19**, 504–516 (2016).
60. Mosher, K. I. & Wyss-Coray, T. Microglial dysfunction in brain aging and Alzheimer's disease. *Biochem. Pharmacol.* **88**, 594–604 (2014).
61. Lalli, M. A. *et al.* Whole-genome sequencing suggests a chemokine gene cluster that modifies age at onset in familial Alzheimer's disease. *Mol. Psychiatry* **20**, 1294–1300 (2015).

62. Heneka, M. T. *et al.* Neuroinflammation in Alzheimer's disease. *Lancet Neurol.* **14**, 388–405 (2015).
63. Lucin, K. M. & Wyss-Coray, T. Immune activation in brain aging and neurodegeneration: too much or too little? *Neuron* **64**, 110–122 (2009).
64. McGeer, P. L., McGeer, E., Rogers, J. & Sibley, J. Anti-inflammatory drugs and Alzheimer disease. *Lancet* **335**, 1037 (1990).
65. Côté, S. *et al.* Nonsteroidal anti-inflammatory drug use and the risk of cognitive impairment and Alzheimer's disease. *Alzheimers Dement.* **8**, 219–226 (2012).
66. Vlad, S. C., Miller, D. R., Kowall, N. W. & Felson, D. T. Protective effects of NSAIDs on the development of Alzheimer disease. *Neurology* **70**, 1672–1677 (2008).
67. Manthripragada, A. D. *et al.* Non-steroidal anti-inflammatory drug use and the risk of Parkinson's disease. *Neuroepidemiology* **36**, 155–161 (2011).
68. Holmes, C., Cunningham, C., Zotova, E., Culliford, D. & Perry, V. H. Proinflammatory cytokines, sickness behavior, and Alzheimer disease. *Neurology* **77**, 212–218 (2011).
69. Holmes, C. *et al.* Systemic inflammation and disease progression in Alzheimer disease. *Neurology* **73**, 768–774 (2009).
70. Zhang, B. *et al.* Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* **153**, 707–720 (2013).
71. Linnartz, B. & Neumann, H. Microglial activatory (immunoreceptor tyrosine-based activation motif)- and inhibitory (immunoreceptor tyrosine-based inhibition motif)-signaling receptors for recognition of the neuronal glycocalyx. *Glia* **61**, 37–46 (2013).
72. Jonsson, T. *et al.* Variant of *TREM2* associated with the risk of Alzheimer's disease. *N. Engl. J. Med.* **368**, 107–116 (2013).
73. Guerreiro, R. *et al.* *TREM2* variants in Alzheimer's disease. *N. Engl. J. Med.* **368**, 117–127 (2013).
74. Naj, A. C. *et al.* Common variants at *MS4A4/MS4A6E*, *CD2AP*, *CD33* and *EPHA1* are associated with late-onset Alzheimer's disease. *Nature Genet.* **43**, 436–441 (2011).
75. Hollingworth, P. *et al.* Common variants at *ABCA7*, *MS4A6A/MS4A4E*, *EPHA1*, *CD33* and *CD2AP* are associated with Alzheimer's disease. *Nature Genet.* **43**, 429–435 (2011).
76. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genet.* **45**, 1452–1458 (2013).
77. Spencer, B. & Masliah, E. Immunotherapy for Alzheimer's disease: past, present and future. *Front. Aging Neurosci.* **6**, 114 (2014).
78. Conboy, I. M. *et al.* Rejuvenation of aged progenitor cells by exposure to a young systemic environment. *Nature* **433**, 760–764 (2005).  
**This landmark study provides the first molecular evidence that heterochronic parabiosis can rejuvenate and reverse stem cell ageing in numerous tissues.**
79. Ruckh, J. M. *et al.* Rejuvenation of regeneration in the aging central nervous system. *Cell Stem Cell* **10**, 96–103 (2012).
80. Villeda, S. A. *et al.* Young blood reverses age-related impairments in cognitive function and synaptic plasticity in mice. *Nature Med.* **20**, 659–663 (2014).  
**The first report to show that systemic administration of plasma from young mice can reverse cognitive deficits in aged mice.**
81. Katsimpardi, L. *et al.* Vascular and neurogenic rejuvenation of the aging mouse brain by young systemic factors. *Science* **344**, 630–634 (2014).
82. Salpeter, S. J. *et al.* Systemic regulation of the age-related decline of pancreatic  $\beta$ -cell replication. *Diabetes* **62**, 2843–2848 (2013).
83. Loffredo, F. S. *et al.* Growth differentiation factor 11 is a circulating factor that reverses age-related cardiac hypertrophy. *Cell* **153**, 828–839 (2013).
84. Baht, G. S. *et al.* Exposure to a youthful circulation rejuvenates bone repair through modulation of  $\beta$ -catenin. *Nature Commun.* **6**, 7131 (2015); erratum **6**, 7761 (2015).
85. Sinha, M. *et al.* Restoring systemic GDF11 levels reverses age-related dysfunction in mouse skeletal muscle. *Science* **344**, 649–652 (2014).
86. Conboy, M. J., Conboy, I. M. & Rando, T. A. Heterochronic parabiosis: historical perspective and methodological considerations for studies of aging and longevity. *Aging Cell* **12**, 525–530 (2013).
87. Castellano, J. M., Kirby, E. D. & Wyss-Coray, T. Blood-borne revitalization of the aged brain. *JAMA Neurol.* **72**, 1191–1194 (2015).
88. Kim, M. J., Miller, C. M., Shadrach, J. L., Wagers, A. J. & Serwold, T. Young, proliferative thymic epithelial cells engraft and function in aging thymuses. *J. Immunol.* **194**, 4784–4795 (2015).
89. Middeldorp, J. *et al.* Preclinical assessment of young blood plasma for Alzheimer disease. *JAMA Neurol.* <http://dx.doi.org/10.1001/jamaneurol.2016.3185> (2016).
90. Sonntag, W. E., Ramsey, M. & Carter, C. S. Growth hormone and insulin-like growth factor-1 (IGF-1) and their influence on cognitive aging. *Ageing Res. Rev.* **4**, 195–212 (2005).
91. Trejo, J. L., Carro, E. & Torres-Aleman, I. Circulating insulin-like growth factor I mediates exercise-induced increases in the number of new neurons in the adult hippocampus. *J. Neurosci.* **21**, 1628–1634 (2001).
92. Stern, S. A., Kohtz, A. S., Pollonini, G. & Alberini, C. M. Enhancement of memories by systemic administration of insulin-like growth factor II. *Neuropsychopharmacol.* **39**, 2179–2190 (2014).
93. Friedlander, A. L. *et al.* One year of insulin-like growth factor I treatment does not affect bone density, body composition, or psychological measures in postmenopausal women. *J. Clin. Endocrinol. Metab.* **86**, 1496–1503 (2001).
94. Baker, L. D. *et al.* Effects of growth hormone-releasing hormone on cognitive function in adults with mild cognitive impairment and healthy older adults. *Arch. Neurol.* **69**, 1420–1429 (2012).
95. Zhang, G. *et al.* Hypothalamic programming of systemic ageing involving IKK- $\beta$ , NF- $\kappa$ B and GnRH. *Nature* **497**, 211–216 (2013).  
**Provides genetic evidence that the hypothalamus controls age-related inflammatory changes in the periphery.**
96. Walker, R. G. *et al.* Biochemistry and biology of GDF11 and myostatin: similarities, differences, and questions for future investigation. *Circ. Res.* **118**, 1125–1141 (2016).
97. Schafer, M. J. *et al.* Quantification of GDF11 and myostatin in human aging and cardiovascular disease. *Cell Metab.* **23**, 1207–1215 (2016).

**Acknowledgements** I would like to thank T. Montine at Stanford University for his critical reading of the manuscript. This work was supported by the US Department of Veterans Affairs and the US National Institute on Aging (AG045034).

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The author declares competing financial interests: see [go.nature.com/2ekr43u](http://go.nature.com/2ekr43u). Readers are welcome to comment on the online version of this paper at [go.nature.com/2ekr43u](http://go.nature.com/2ekr43u). Correspondence should be addressed to T.W.-C. ([twc@stanford.edu](mailto:twc@stanford.edu)).

**Reviewer Information** *Nature* thanks F. Gage, M. Mattson and the other anonymous reviewer(s) for their contribution to the peer review of this work.



# The road to restoring neural circuits for the treatment of Alzheimer's disease

Rebecca G. Canter<sup>1</sup>, Jay Penney<sup>1</sup> & Li-Huei Tsai<sup>1,2</sup>

**Alzheimer's disease is a progressive loss of memory and cognition, for which there is no cure. Although genetic studies initially suggested a primary role for amyloid- $\beta$  in Alzheimer's disease, treatment strategies targeted at reducing amyloid- $\beta$  have failed to reverse cognitive symptoms. These clinical findings suggest that cognitive decline is the result of a complex pathophysiology and that targeting amyloid- $\beta$  alone may not be sufficient to treat Alzheimer's disease. Instead, a broad outlook on neural-circuit-damaging processes may yield insights into new therapeutic strategies for curing memory loss in the disease.**

Age-related dementias will affect almost 10% of people in the United States, and these conditions place a tremendous burden on such individuals and their families<sup>1</sup>. The most prevalent type of dementia, Alzheimer's disease (AD), causes a devastating and progressive loss of cognition, for which there is no effective treatment or cure. Analyses of the brains of people with AD suggest that the presence of extracellular aggregates of amyloid- $\beta$  peptides, intracellular inclusions of neurofibrillary tangles rich in microtubule-associated protein tau and neuritic plaques are pathological hallmarks of the disease, yet there is no conclusive link between these observations and the cognitive symptoms<sup>2</sup>. The inability to definitively connect progressive memory loss to biomarkers greatly impedes the quest for effective therapeutic interventions for AD, but enhanced efforts to understand mechanisms of cognitive decline are revealing new avenues for intervention.

The insidious onset of AD-related memory loss has hampered progress towards effective therapies because cognitive symptoms emerge late in the progression of the disease. At this stage, there is already extensive deposition of amyloid- $\beta$ , formation of neurofibrillary tangles and cell death, each of which might contribute to memory loss<sup>2</sup>. The first genetic studies performed in the 1990s indicated that amyloid- $\beta$  was the main causative factor in AD because mutations or duplications in the gene *APP*<sup>3</sup>, which encodes the amyloid- $\beta$  precursor protein (APP, also known as amyloid- $\beta$  A4 protein), and in the genes *PSEN1* and *PSEN2*, which encode amyloid- $\beta$  processing-pathway components presenilin-1 and presenilin-2 (refs 4 and 5), lead to inherited, early onset AD. This breakthrough in understanding laid the foundation for the amyloid cascade hypothesis, which postulates that the accumulation of amyloid- $\beta$  peptides initiates the pathophysiology of AD, leading to neurofibrillary tangles and neurodegeneration that cause memory loss<sup>6</sup>. As a corollary to this hypothesis, increased levels of amyloid- $\beta$  in the brain — a result of the excessive generation of peptides or an inability to clear them — underlie cognitive dysfunction, and treatments to reduce the load of amyloid- $\beta$  should therefore improve cognition<sup>7</sup>. Despite extensive evidence to support amyloid- $\beta$ -driven neurodegeneration, drugs targeted at amyloid- $\beta$  have failed to reverse deficits in memory or to halt cognitive decline, even alongside a considerable reduction in amyloid- $\beta$ . As treatment strategies that are based on amyloid- $\beta$  modification shift to preventing the accumulation of amyloid- $\beta$  peptides, the refinement of drug targets and improving compound design, advances in our understanding of how brain functions are disrupted in AD, coupled with data from omics technologies and brain-imaging platforms, are highlighting other promising

routes towards ameliorating cognitive decline.

As well as the genes in the amyloid- $\beta$  pathway, other human genomic loci have been identified that increase the risk of developing AD considerably. Although some of these genes have been linked to amyloid metabolism, many appear to function in cell signalling pathways and in the immune system of the brain<sup>8</sup>, which indicates the importance of atypical cellular responses to changing brain states in AD. Emerging functional data are also shifting the aetiological focus of the disease from a neuron-centric view to an integrated outlook that acknowledges the synergistic functions of the different cell types of the brain<sup>9</sup>. Consistent with this holistic view, numerous disrupted processes are interconnected in AD and interact to provoke a cycle of dysfunction as the disease progresses<sup>9</sup>. This array of altered cellular processes ultimately disrupts neural circuits and network connectivity<sup>10</sup>, which can trigger further coping mechanisms in cells and might also propagate the spread of AD-related protein or peptide aggregates.

The complexity of these interactions makes disentangling the causes of AD a daunting task. Although systematic examination of the genetic and cellular changes can provide mechanistic insights into the cascade of events that occur during AD<sup>9</sup>, only understanding how these alterations contribute to the progressive dysfunction of brain networks and circuits will connect direct observations in cells to biologically abstract cognitive impairment. The identification of vulnerable networks and susceptibility nodes within them might also provide clues to the origins and progression of AD. This Review will explore briefly the role of amyloid- $\beta$  in the development of AD. Rooted in newly identified genetic risk factors, it will then examine growing evidence that suggests that genetic, cellular and circuit dysregulation results from, and can lead to, the accumulation of amyloid- $\beta$  peptides, the aggregation of tau proteins and cognitive impairment. With such a perspective, the AD field can move from cognitively unsuccessful clinical trials targeted at amyloid- $\beta$  towards multipronged approaches to treatment that incorporate therapies that target cells and circuits.

## AD through the lens of the amyloid cascade hypothesis

The accumulation of amyloid- $\beta$  plays an important part in the pathophysiology of AD. Although the physiological functions of amyloid- $\beta$  remain unknown, decades of research founded in genetic studies of AD continue to suggest a causative role for the peptide. The build-up of amyloid- $\beta$  in the brain parenchyma probably contributes to the loss of synapses, neurodegeneration and alterations in neuronal activity. Each

<sup>1</sup>The Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. <sup>2</sup>The Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA.

## BOX 1

# The production of amyloid- $\beta$

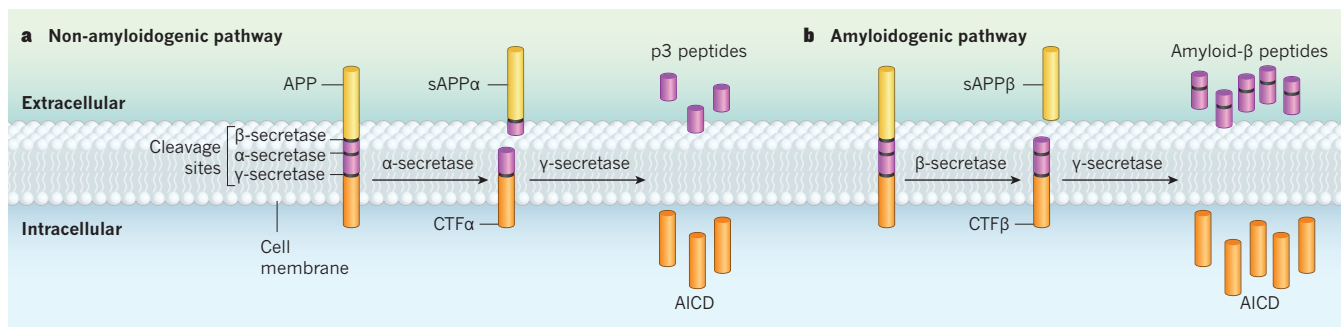
Amyloid- $\beta$  peptides result from the sequential cleavage of APP, a type I integral membrane glycoprotein. The structure of APP includes an amyloid- $\beta$  domain (purple) with cleavage sites for secretase enzymes. Two pathways compete for the APP substrate, which leads to either amyloidogenic or non-amyloidogenic processing of the protein<sup>147</sup> (Box Fig.).

In the non-amyloidogenic pathway (Box Fig. a), the cleavage of APP in the amyloid- $\beta$  domain by  $\alpha$ -secretase, a complex that contains ADAM metalloproteases, releases the soluble ectodomain sAPP $\alpha$  and the C-terminal fragment CTF $\alpha$ . The subsequent cleavage of CTF $\alpha$  by  $\gamma$ -secretase produces a soluble extracellular p3 peptide and the APP intracellular domain (AICD)<sup>148</sup>.

The amyloidogenic pathway (Box Fig. b) involves APP cleavage

by  $\beta$ -secretase (also known as BACE1), which releases the soluble ectodomain sAPP $\beta$  and CTF $\beta$ . Cleavage of CTF $\beta$  by  $\gamma$ -secretase yields amyloid- $\beta$  peptides of varying lengths as well as the AICD fragment. The pathogenic impact of amyloid- $\beta$  peptides varies with length. The longer amyloid- $\beta$ (1–42) and amyloid- $\beta$ (1–43) species are more prone to aggregation and prion-like seeding, and because these structures seem to be more toxic than shorter amyloid- $\beta$  peptides, the ratio of amyloid- $\beta$ (1–42) and amyloid- $\beta$ (1–40) might predict the severity of AD<sup>149</sup>.

New pathways for APP processing are emerging. For example, the  $\eta$ -secretase pathway yields amyloid- $\eta$  fragments that inhibit neuronal function. N-APP, a mediator of axonal degeneration, arises from sAPP $\beta$  following cleavage by  $\beta$ -secretase. And the caspase-mediated cleavage of the AICD can result in transcriptionally active products<sup>147,150</sup>.



of these changes disrupts neural circuits, which can lead to widespread network dysfunction and cognitive decline.

## Familial genetics implicate amyloid- $\beta$ in AD

The most compelling evidence to support a central role for amyloid- $\beta$  in AD comes from studies of the familial form of the disease. Cases of familial AD account for 1–5% (ref. 11) of patients with AD and they are the result of disease-causing, autosomal dominant mutations or duplications in *APP*, or mutations in *PSEN1* and *PSEN2* (refs 3–5). Mutations in *APP*, *PSEN1* and *PSEN2* shift the processing of APP (Box 1) to the amyloidogenic pathway and bias APP cleavage towards the longer, toxic amyloid- $\beta$  peptides<sup>12</sup>. Although the mechanism by which the duplication of *APP* leads to toxicity is less clear, familial-AD-related *APP* duplications exist<sup>13</sup>, and people with Down's syndrome, who have three copies of *APP*, also develop AD<sup>14</sup>. Therefore, despite differences in their type, location in a particular gene or in the genes that they affect, heritable changes that result in the production of longer amyloid- $\beta$  species cause AD. (Mutations in the gene *MAPT*, which encodes tau protein, usually lead to frontotemporal dementia — a similar yet distinct disease<sup>15</sup>.) Amyloid- $\beta$  is implicated further in AD by mutations in *APP* that do not increase levels of amyloid- $\beta$  and do not seem to be pathogenic; instead, these mutations might confer protection from AD. For example, an *APP* variant in the population of Iceland reduces levels of amyloidogenic processing and protects against AD and normal age-related cognitive decline<sup>16</sup>.

As well as mutations that cause familial AD, other heritable genetic risk factors contribute to an individual's susceptibility to late-onset AD and also to amyloid- $\beta$  accumulation. The most important mutation in this context, and the first to be identified, is the  $\epsilon 4$  allelic variant of the gene *APOE* (*APOE4*), which encodes apolipoprotein E (APOE). The presence of *APOE4* leads to a three- or four-fold increase in the likelihood of developing AD<sup>17</sup>, and people with AD who carry *APOE4* show more profound evidence of amyloid- $\beta$  aggregates than do non-carriers with AD<sup>18</sup>. Studies in mice suggest that ApoE regulates amyloid- $\beta$  levels in an ApoE-isoform-dependent manner, such that the  $\epsilon 4$  isoform promotes

amyloid build-up whereas the  $\epsilon 2$  isoform seems to enhance its clearance<sup>19</sup>. Because variants of both *APP* and *APOE* can either increase amyloid- $\beta$  accumulation, which increases susceptibility to AD, or reduce its accumulation, which reduces susceptibility, the genetic data strongly support an important aetiological role for amyloid- $\beta$  accumulation in AD-related cognitive decline.

## Amyloid- $\beta$ impairs synapses and destabilizes circuits

Amyloid- $\beta$  peptides can exert numerous detrimental effects on neurons and the other cell types of the brain<sup>9</sup> (Fig. 1). As a peptide species with acute toxicity, oligomeric amyloid- $\beta$  directly incites neuronal apoptosis through interactions with cell-surface receptors. Moreover, longer-term accumulation of toxic amyloid- $\beta$  species in the parenchyma also leads to oxidative damage of DNA and proteins, to physical injury of cellular organelles and to dysregulation of intracellular calcium levels — each of which can provoke cell death<sup>20</sup>. Under *in vitro* conditions, exposure to amyloid- $\beta$  induces neuronal dysfunction and can cause cell death within hours<sup>21</sup>. Although exposure to amyloid- $\beta$  under *in vitro* conditions can induce neuronal dysfunction and cause cell death within hours, it takes years for amyloid accumulation to have detectable consequences *in vivo*. Despite the difference in timescale, the brains of people with AD show alterations similar to those seen *in vitro*. Most notably, considerable neurodegeneration accompanies the onset of cognitive decline<sup>22</sup>. However, amyloid- $\beta$ -induced deficits in synaptic plasticity, circuit function and cognition might develop before this cell loss occurs<sup>10</sup>. Indeed, this possibility is supported by the observation that the levels of amyloid- $\beta$  in people with AD seem to plateau before the onset of rapid neurodegeneration and cognitive symptoms<sup>22</sup>, and *ex vivo* studies in rats demonstrate that synapse loss occurs in response to small elevations in amyloid- $\beta$  levels<sup>23</sup>. In people with AD, it is therefore possible that synapse loss precedes the formation of overt aggregates, and the loss of these synapses has been shown to contribute to circuit dysfunction and cognitive decline in mouse models of AD<sup>24,25</sup>.

There is abundant evidence to support a physiological role for



amyloid- $\beta$  at the synapse, and the data also suggest that synaptic activity is an important source of amyloid- $\beta$  in the parenchyma<sup>26</sup>. Excitatory activity promotes the proteolysis of APP and its release into the extracellular space<sup>27–29</sup>. The effects of amyloid- $\beta$  on synapse activity vary with its extracellular concentration: low levels of amyloid- $\beta$  promote excitatory activity and higher levels depress it. Small increases in amyloid- $\beta$  levels promote activity through presynaptic acetylcholine receptors, which elevate internal calcium concentrations to increase the probability of glutamate release<sup>30,31</sup>. Although the postsynaptic excitation that results could lead to positive feedback in which the further release of amyloid- $\beta$  increases synaptic excitability, increasingly high levels of amyloid- $\beta$  actually depress synapse activity through several mechanisms that modify synapse strength, including the internalization of glutamate receptors<sup>32,33</sup>. Whereas acute increases in synaptic amyloid- $\beta$  can impair the long-term potentiation of synaptic strength and promote the depression of synaptic activity (known as long-term depression)<sup>34–36</sup>, chronic elevations can weaken connectivity, alter the dynamics of dendritic spines, promote synapse loss and impair circuit-shaping processes that underlie learning and memory<sup>27,34</sup>. Amyloid- $\beta$ -induced loss of dendritic spines can lead to hyperexcitable neurons that fire more easily<sup>24</sup>, and amyloid- $\beta$  also alters the balance between excitatory and inhibitory activity by influencing inhibitory interneurons. Loss of synaptic inhibition occurs through numerous pathways, including the downregulation of cell-surface voltage-gated sodium channels. Indeed, reduced expression of the Na<sub>v</sub>1.1 sodium channel subunit (SCN1A) in a mouse model of AD hinders the propagation of action potentials through inhibitory parvalbumin-expressing interneurons, which results in a reduction in the release of the primary inhibitory neurotransmitter, GABA ( $\gamma$ -aminobutyric acid) and the loss of inhibition on excitatory neurons<sup>37</sup>.

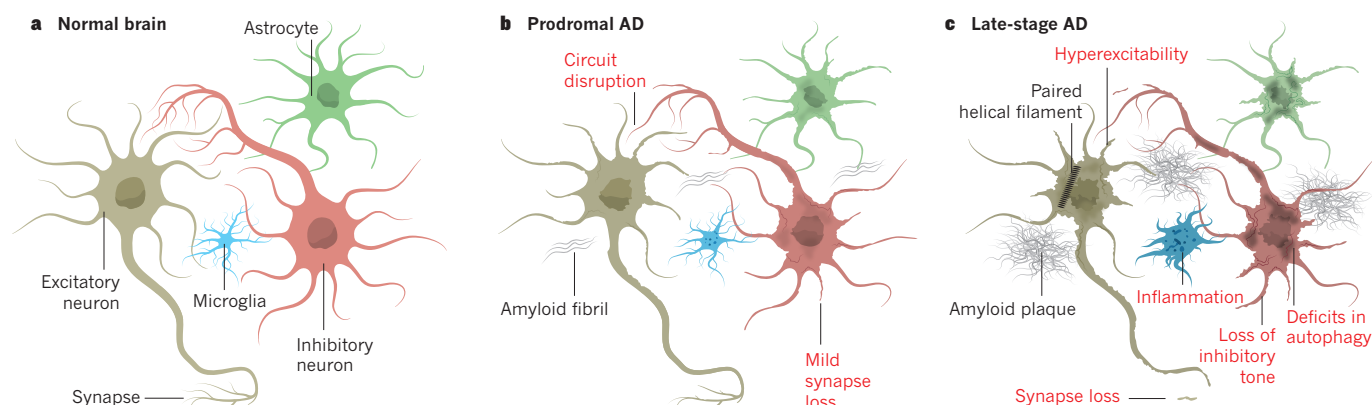
Long-term amyloid- $\beta$  accumulation, disinhibition of excitatory cells and synaptic loss lead to neuronal hyperactivity, which occurs in brain regions associated with learning and memory (such as the hippocampus) in some presymptomatic individuals<sup>38</sup>. In time, this can lead to epileptiform activity. In a mouse model of AD-like amyloid- $\beta$  accumulation, such activity induces the compensatory sprouting of inhibitory axons that can impair learning processes<sup>25</sup>. People with AD who have seizures exhibit worse cognitive outcomes<sup>39</sup>, and the cognition of individuals with mild cognitive impairment is restored temporarily when they are treated with the anti-epileptic drug levetiracetam to reduce abnormal activity<sup>40</sup>. The correlation between altered activity and cognitive performance suggests that these aberrant patterns of activity might be directly related to memory impairment. Counteracting hyperactivity in mouse models of AD not only rescues local circuit dynamics<sup>41</sup> but

also reinstates long-range network coherence to repair learning<sup>42</sup>.

### Network susceptibility is crucial for amyloid- $\beta$ driven AD

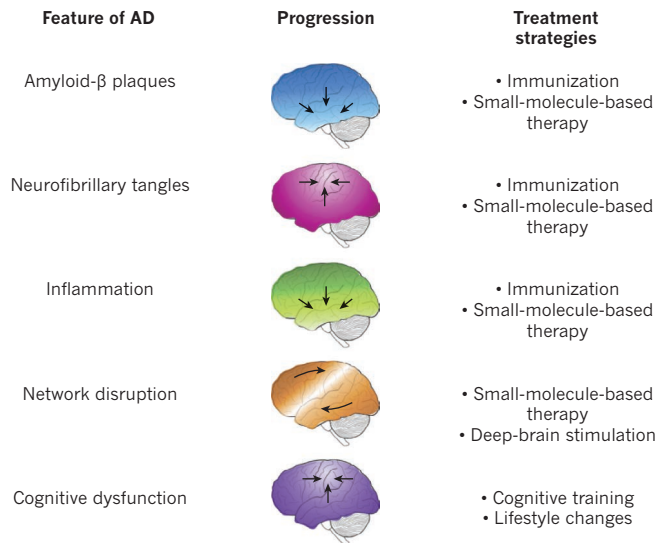
The deposition of amyloids affects circuit connectivity and network activity. Therefore, identification of the brain networks that are susceptible to amyloid- $\beta$ -induced dysfunction should reveal the pattern of disease propagation through the brain and could explain how the spread of amyloid- $\beta$  contributes to the symptoms of AD. Initial cross-sectional studies of post-mortem tissue were used to create a hierarchical map of the advance of amyloid- $\beta$  through the brain<sup>43</sup>. This led to the observation that amyloid- $\beta$  deposition that begins in neocortical regions involved in cognition spreads to the neural hubs that underlie learning and memory, and finally progresses to the motor and sensory structures. It also provided a framework for understanding the successive deterioration of cognitive and sensory impairments<sup>43</sup>. Although the sequential involvement of interconnected brain regions was supported by these findings, post-mortem studies of amyloids in the brain could not provide insights into the network pattern or mechanisms of its spread, or correlate low levels of amyloid deposition with a definitive diagnosis of AD in people who died before developing memory impairment. The need to understand the processes that underlie AD progression, and to connect the spread of amyloid to the development of AD, therefore paved the way for longitudinal studies of amyloid- $\beta$  accumulation in people with the disease.

The development of tracers for positron emission tomography (PET) that bind amyloid- $\beta$ , such as <sup>11</sup>C-Pittsburgh compound B (PIB) and <sup>18</sup>F-florbetapir, has enabled the pattern of amyloid- $\beta$  deposition to be assessed across the course of AD<sup>44,45</sup> (Fig. 2). Cross-sectional PET studies of the brains of people with AD confirm neuropathological findings that show considerable deposits of amyloid- $\beta$  throughout the neocortex and also demonstrate that the amount of cortical retention is predictive of cognitive decline<sup>46,47</sup>. These findings support the genetic data and indicate that amyloid- $\beta$  levels might be a good biomarker for AD<sup>48</sup>. Consistency in the locations of tracer retention in PET studies suggests that the accumulation of amyloid- $\beta$  in an AD-specific pattern is a fundamental change that affects cognitive outcomes, and memory decline could be related to the deposition of amyloid- $\beta$  in regions of the default mode network (DMN)<sup>49</sup>. The DMN is a group of functionally connected regions of the brain that are co-activated during passive thinking, remembering and planning. Initial investigations using PIB-PET show an overlap between DMN regions and areas with high levels of amyloid- $\beta$ , which suggests that the DMN might be vulnerable to pathophysiological processes in AD<sup>50</sup>. DMN susceptibility has been demonstrated by functional connectivity magnetic resonance imaging (MRI) studies that show reduced connectivity within



**Figure 1 | Neural circuits and synapses during the progression of AD.** **a**, Subpopulations of neurons and glial cells form functional circuits through synaptic connections. **b**, In prodromal AD, amyloid- $\beta$  fibrils begin to form in the extracellular space, possibly contributing to early circuit dysfunction that stimulates inhibitory sprouting and the initiation of inflammatory processes. **c**, In late-stage AD, amyloid- $\beta$  plaques grow as production of the peptide outpaces its clearance. Intracellular neurofibrillary tangles consisting of paired helical

filaments of hyperphosphorylated tau protein also form. Deficits in autophagy and other quality-control pathways contribute to the dysfunction of neurons and glia. The activation and proliferation of glial cells promotes inflammation and can affect circuit function in numerous ways. Several mechanisms affect GABAergic signalling and contribute to the loss of inhibitory tone. Despite the loss of synapses, excitatory neurons become hyperexcitable, compromising the fidelity of synaptic network connections for long-range communication.



**Figure 2 | Hallmarks of AD in the brain and corresponding therapeutic strategies.** Distinctive features of AD include amyloid- $\beta$  plaques, neurofibrillary tangles and cognitive dysfunction<sup>2</sup>. Network disruption and inflammation are also important mediators of the AD state<sup>49,146</sup>. Each type of alteration seems to start in a specific functional subregion of the brain and spreads through networks to encompass most of the organ in late-stage AD. The most appropriate current strategies for the treatment of each feature are listed.

the network<sup>49</sup>. About 25% of people with good cognitive health show a high retention of PIB, which correlates with reductions in DMN connectivity and with worse episodic memory and predicts the diagnosis of dementia in follow-up studies<sup>51,52</sup>. Together, this evidence suggests that alterations in the DMN might underlie cognitive impairment. Surprisingly, despite cognitive deterioration, increases in amyloid- $\beta$  levels over the conversion period from the condition amnesic mild cognitive impairment to AD are relatively modest. This is consistent with observations that people with AD show few differences in the retention of PIB as the disease progresses<sup>22,53</sup>. Instead, it is possible that progressive DMN dysfunction contributes to clinical deterioration<sup>54</sup>.

Although imaging studies in humans show that the localization of amyloid- $\beta$  correlates with cortical network dysfunction, these techniques have yet to reveal initiating patterns of amyloid- $\beta$  accumulation or give insights into the mechanisms of amyloid- $\beta$  propagation. Some preclinical studies suggest that amyloid- $\beta$  might propagate in a prion-like manner and undergo cell-to-cell transmission<sup>55,56</sup>. However, other studies propose that amyloid- $\beta$ -induced circuit dysfunction affects network connectivity<sup>42</sup> and that local aberrant activity could lead to the accumulation of amyloid- $\beta$  at downstream projection structures, contributing to the sequential appearance of amyloid- $\beta$  in regions of connected networks. Whether functional or physical mechanisms contribute to amyloid- $\beta$  progression, the large temporal dissociation between the accumulation of amyloids and the emergence of overt memory loss underscores the urgency of identifying regions that show susceptibility to amyloid- $\beta$  deposition for the early detection of and intervention in AD.

### Amyloid- $\beta$ -based treatments have yet to restore cognition

Despite various complexities, it is clear that amyloid- $\beta$  contributes to the progression of AD, which makes it a prime target for therapeutic intervention (Fig. 2). Because elevated levels of amyloid- $\beta$  probably underlie its pathogenicity, treatment strategies have emerged that target the two determinants of amyloid- $\beta$  load: the aberrant generation of amyloid- $\beta$  and faulty clearance mechanisms.

Drugs that modulate the activity of the enzymes  $\beta$ -secretase or  $\gamma$ -secretase limit amyloidogenic proteolysis and are the main therapeutic strategy for reducing amyloid- $\beta$  production. Phase II clinical trials have identified several safe first-generation drugs for each enzyme<sup>57,58</sup>.

However, most such trials have failed because of problems with target specificity, brain permeability or study design without testing cognitive or amyloid- $\beta$  outcomes<sup>59,60</sup>. The few trials that reported main outcomes showed reductions in markers of amyloid- $\beta$  in cerebrospinal fluid but with conflicting cognitive results, including hastened decline<sup>61</sup>. Given these pharmacological setbacks and uninformative outcomes with respect to memory, it is difficult to make conclusions about the feasibility of using proteolytic enzymes as targets for AD treatment. However, redesigned clinical trials with updated dosing and outcome design and an improved understanding of the functions of the enzyme targets should help to ensure the safety and effectiveness of the next generation of compounds.

Enhancing amyloid- $\beta$  clearance from the parenchyma is an alternative approach to modifying amyloid- $\beta$  levels, and this is being pursued mainly through active and passive strategies of immunization against the peptide. Initial clinical trials that investigated active immunization reduced the amyloid- $\beta$  burden successfully but also caused severe side effects<sup>62</sup>. The more positive effects on amyloid- $\beta$  levels have facilitated the development of a new generation of immunotherapies that have advanced quickly through clinical trials. In general, people with AD tolerate these newer immunotherapies well, and phase I and phase II trials have demonstrated considerable reductions in amyloid- $\beta$  levels<sup>63</sup>. These results are tempered by cognitive performance outcomes that suggest that such treatments might slow decline only in prodromal or mild AD. Furthermore, the efficacy of amyloid- $\beta$  immunization strategies in this population is not consistent across clinical trials<sup>64,65</sup>. A more successful study used florbetapir-based PET imaging to identify people with both AD and amyloid- $\beta$  deposits<sup>66</sup>. People who lack such deposits are unlikely to respond to amyloid- $\beta$  targeted therapies. By choosing participants with AD and amyloid- $\beta$  deposits, the study was therefore able to test an amyloid- $\beta$  targeted antibody therapy in individuals who are most likely to respond. The positive findings indicated that the stratification of people with AD, and the identification of biomarkers that correlate with those groups, might be crucial for testing treatments with different mechanisms of action<sup>66</sup>. As clinical trials advance through phase III and phase IV, further insights into the lasting cognitive outcomes could provide an improved understanding of the efficacy of these strategies for long-term prognosis.

The results from numerous amyloid- $\beta$ -modifying clinical trials herald considerable progress towards reducing the amyloid- $\beta$  burden. Although they are varied and minor, the cognitive benefits of such treatments suggest that decreasing amyloid- $\beta$  levels slows the progression of AD. However, because both cognitive effects and the safe removal of amyloid are seen in the initial stages of the disease, and because many patients first seek treatment after they experience memory loss, preventing further deterioration is only the first step in an effective treatment regimen. The identification of additional interventions that can restore cognitive function and treat cellular and circuit dysfunction at later stages of the disease will be crucial for improving the quality of life of people with AD and to avert the looming public-health crisis.

### Numerous factors induce network dysfunction in AD

The accumulation of amyloid- $\beta$  in the brain has deleterious effects, and people with familial AD and rodent models of AD demonstrate considerable pathological and cognitive alterations in response to the peptide's presence. However, most people with AD have considerable accumulation of amyloid- $\beta$  without causative mutations in genes related to APP processing. With the genetic evidence mounting for the contribution of other disrupted cellular mechanisms to the development of AD, understanding how these processes interact with amyloid- $\beta$  aggregates, contribute to neurodegeneration and affect circuit function to produce cognitive decline is important for implementing effective treatment strategies to not only slow, but to also halt and reverse, memory loss.

### Genetic risk and dysregulation prime the brain for AD

Cases of AD with no identifiable genetic cause account for more than 95% of individuals who develop symptoms in later life<sup>67</sup>. Similar to those with



familial AD, people with late-onset AD experience amnesic memory loss and develop amyloid- $\beta$  aggregates and neurofibrillary tangles; however, the levels of disease in late-onset AD vary considerably<sup>2</sup>. Because variation between individuals is high, and because numerous complex and inter-related processes can lead to amyloid- $\beta$  accumulation and tau aggregation, understanding the mechanisms that drive late-onset AD has proven more difficult than disentangling familial AD<sup>9</sup>. Although functional studies of the amyloid- $\beta$  peptide in models of familial AD have yielded insights into the mechanisms of peptide accumulation, the description of further intracellular functions for APP and the presenilins, together with the identification of increasing numbers of genetic risk alleles and regulatory mechanisms, has shed light on other processes disrupted in AD that might lead to amyloid- $\beta$  accumulation and neurofibrillary tangle formation and contribute to cognitive decline.

Although protein changes related to familial AD increase amyloid- $\beta$  levels directly, they might also have consequences for processes other than APP proteolysis. Mutations in *APP*, *PSEN1* and *PSEN2* could affect other roles of the proteins that these genes encode, including cholesterol binding, cell adhesion, cytoskeletal dynamics, ion homeostasis, endocytosis and synaptic plasticity<sup>68,69</sup>, all of which can affect the properties of circuits. Mice that lack APP have deficits in the induction of long-term potentiation and also show age-dependent cognitive decline<sup>70,71</sup>. Although familial AD-linked mutations in the presenilins have a clear effect on amyloid- $\beta$  production, it is possible that the disruption of the other physiological functions of the presenilins also contributes to circuit dysfunction and memory impairment. For example, deletion of the presenilins suggests that they have crucial roles in neural development, synaptic plasticity and memory performance, and mutations in familial AD affect the function of these proteins as calcium-leak channels in the endoplasmic reticulum<sup>72,73</sup>. Similarly, the association of APOE with the clearance of amyloid- $\beta$  drew focus away from its roles in regulating cholesterol and lipids; however, studies of the brains of people with AD suggest that alterations in these processes could be an important mechanism of pathogenesis<sup>74</sup>. The identification of new genetic risk factors in recent years has underscored the potential relevance of these processes to cognitive decline.

The increasing accessibility of genome-level analyses has enabled large-scale screens of individuals with AD to identify further genetic risk loci that might increase susceptibility to the disease. These studies have implicated genes that are involved in lipid metabolism and cholesterol homeostasis, cell adhesion, cytoskeletal dynamics, ion regulation, vascularization and maintenance, transcriptional regulation, inflammation and endocytosis (for example, *CLU*, *PICALM*, *SLC24A4*, *TREM2* and *BIN1*) (ref. 8). So far, confirmation of a subset of the newly identified loci in functional experiments has demonstrated that lipid processing, endocytosis and inflammation might contribute substantially to the development of AD<sup>75–77</sup>. Unlike *APOE*, *APP*, *PSEN1* and *PSEN2*, the disease-causing genetic variants and the cellular consequences of alterations at these newly identified risk loci are largely unknown<sup>8</sup>. Although many of the implicated proteins seem to play parts in amyloid- $\beta$  processing or clearance<sup>8</sup>, they might also be important for molecular homeostasis, cell function and synaptic dynamics, all of which influence circuit activity and network connectivity<sup>9,10</sup>.

Similar to encoded genetic risk factors, epigenetic mechanisms can also alter gene expression. People with late-onset AD show altered DNA methylation at several AD genetic risk loci<sup>78</sup>. Association of the methylation status at some of these loci with the risk of AD and with cellular processes that are linked to the disease also indicates a functional significance for such alterations<sup>79</sup>. Studies in a mouse model of AD confirm the potential for epigenetic contributions to the disease state and suggest that the progression of the disease is associated with widespread shifts in histone modifications in hippocampal cells, which correlates with the transcriptional repression of synaptic genes and an induction of immune-related genes<sup>80</sup>. Evolutionarily conserved immune-gene regulatory regions with altered histone modification profiles are enriched for genetic variants identified in people with AD, which supports the genetic links between immune processes and the risk of AD<sup>80</sup>. Further insights

into the epigenetic changes that are associated with neuronal genes in the disease comes from both rodent models and studies of people with AD, both of which have elevated levels of the enzyme histone deacetylase 2, an important negative regulator of synaptic plasticity<sup>81</sup>.

Other factors, such as circuit activity and peripheral signalling, can also modulate genetic risk by acting on transcriptional components, which could lead to fundamental changes in cellular health. An example of a protective transcription factor is RE1-silencing transcription factor (REST), which upregulates cellular protection pathways in ageing<sup>82</sup>. Higher levels of REST expression are correlated with increased hippocampal volume, improved cognitive outcomes with age and protection against amyloid- $\beta$  toxicity. Levels of REST are lower in people with AD, and although the mechanism that underlies these reduced levels remains unclear, the time-dependent expression kinetics of REST could provide clues to the influence of circuits in healthy ageing. REST is regulated by non-cell-autonomous signalling, which indicates the vital role of neuronal communication and circuit function in neuroprotection against AD<sup>82</sup>.

### Molecular alterations affect cellular health

Genetic loci identified in the past 5 years through enhanced screening of people with AD have helped to pinpoint cellular processes that might be involved in the development of AD and highlight the importance of both neurons and non-neuronal cells in disease progression<sup>9</sup> (Fig. 1). There is considerable evidence to support the idea that dysfunction across different cell types in the brain can lead to circuit dysregulation and cognitive impairment. The cellular phase of AD refers to the time during which processes that normally contribute to feedback and feed-forward loops go awry, resulting in cellular disease states. This progression is aligned with the observation that the deposition of amyloid- $\beta$  plateaus before the onset of memory symptoms<sup>9</sup>. Although cognition remains largely intact at the initiation of deposition, it is probable that progressive changes during and following the cellular phase of AD alter network connections and circuit activity to produce cognitive decline<sup>10</sup>. Because the events that lead to amyloid- $\beta$  aggregation in the absence of causative mutations remain vague, understanding the convergent cellular factors that might affect circuit activity, drive cognitive deterioration and initiate amyloid- $\beta$  accumulation is crucial for restoring memory capacity in advanced AD.

A number of cellular processes that are implicated by AD genetic risk loci might contribute to the onset and progression of AD when aberrant. These include mitochondrial function, the oxidative stress response, DNA damage repair and calcium homeostasis. Although each locus has not yet been confirmed, molecular studies imply that many such processes go awry in AD. Mitochondria accumulate age-related damage that can decrease efficiency, increase the release of reactive oxygen species, and contribute to other changes, including the oxidation of proteins and lipids and  $\beta$ -secretase (also known as BACE1) upregulation<sup>83,84</sup>. Mitochondria are the main producers of reactive oxygen species, the release of which can cause the relaxation of heterochromatin, a tightly packaged form of DNA that represses certain portions of the genome<sup>85</sup>. The brains of people with AD show reduced heterochromatin-associated methylation marks, as do animals that overexpress tau protein, and this change is correlated with the expression of genes that are normally repressed in healthy individuals<sup>85</sup>. Detectable breaks in DNA occur at the initial stages of disease progression in mouse models of AD and in the brains of people with the disease<sup>86,87</sup>. These breaks might be the result of physiological processes: DNA breaks arise in response to neuronal activity, although amyloid- $\beta$  can exacerbate the damage<sup>88</sup>. Because the breaks that form during learning-related activity initiate the transcription of genes that are important for learning and memory, it is possible that as DNA repair mechanisms fail with age, mutations and breaks accumulate at plasticity-related loci, which stunts cognition<sup>89,90</sup>. This suggests a direct link between neural circuit activity, DNA integrity and memory.

Calcium is a crucial mediator of many intracellular events and calcium levels regulate synaptic plasticity and signal propagation<sup>91</sup>. Brain cells regulate cytosolic calcium levels tightly via calcium influx through cell-surface channels and some neurotransmitter receptors and storing

calcium in the endoplasmic reticulum<sup>91</sup>. In AD, a number of processes increase cytosolic calcium, including damage to mitochondrial or plasma membranes, circuit hyperactivity and amyloid- $\beta$ -induced calcium influx across the plasma membrane. These events then trigger the release of calcium from the endoplasmic reticulum. Interestingly, PSEN1 has been shown to be an important regulator of such calcium release<sup>73</sup>. The AD risk gene *SLC24A4* encodes a potassium-dependent sodium–calcium exchanger and mutations could alter cellular calcium handling and increase the risk of developing late-onset AD<sup>8</sup>. Furthermore, calcium can induce signalling cascades; for example, the calcium-mediated cleavage of the cyclin-dependent kinase 5 (Cdk5) activator p35 to p25 by calpain regulates synaptic depression in neurons under physiological conditions<sup>92</sup>. However, under pathogenic conditions, enhanced p25–Cdk5 signalling results in increases in histone deacetylase 2, synaptic depression, dendritic spine loss, neurodegeneration, inflammation and aberrant tau phosphorylation, which correlates with the presence of neurofibrillary tangle-like aggregates in mouse models of AD<sup>80,81,92</sup>. Elevated calcium levels in the cell can also produce ionic imbalance, facilitate the aberrant release of presynaptic neurotransmitters and dysregulate postsynaptic signal transduction that ultimately alters gene regulation, cellular dynamics and neuronal activity<sup>31,91,92</sup>.

### Intracellular pathophysiology impairs circuits

Neuronal activity and circuit connectivity rely on connections between cells that can span long distances. This specialized morphology necessitates highly coordinated intracellular functions, however, AD disrupts processes that maintain neuronal health and support circuit function. The hallmark intracellular deposits of AD, neurofibrillary tangles, are the result of aberrant aggregation of tau protein, which usually associates with microtubules. Microtubules are the primary cytoskeletal component that gives neurons structure and enables trafficking. The stability and functionality of microtubule assemblies depends partially on the dynamic binding of tau<sup>93,94</sup>. Aggregates of tau form when hyperphosphorylated tau proteins assemble into paired helical filaments and the presence of these structures is a requisite for the diagnosis of AD<sup>2</sup>. Furthermore, the level of neurofibrillary tangles correlates strongly with the progression of AD and cognitive symptoms<sup>2</sup>. Tau binds microtubules in its unphosphorylated state, and site-specific phosphorylation weakens tau's microtubule-binding capacity, which releases tau into the cytosol and reduces its solubility<sup>93,94</sup>. Dephosphorylation by protein phosphatase 2A (PP2A) maintains appropriate tau dynamics, which enables tau to actively bind and be released from the cytoskeletal network<sup>95,96</sup>. But in AD, circuit hyperactivity and calcium influx trigger aberrant tau phosphorylation by kinases such as Cdk5, mitogen-activated protein kinase (MAPK) and protein kinase A (PKA) and induce concomitant dysfunction of PP2A that leads to tau hyperphosphorylation, insolubility and aggregation<sup>95,97</sup>. Although the pathogenicity of tau was assumed to be related to its aggregation and apparent mislocalization, tau also has an important role at the synapse, where it may coordinate the postsynaptic density and, similar to amyloid- $\beta$ , may be released in an activity-dependent manner<sup>98–100</sup>. Further observations of the aberrant acetylation of tau and the accumulation of tau-rich autophagosomes in the brains of people with AD suggest that pathways involving tau and certain genetic risk factors (for example, the gene *PICALM*) converge, indicating that tau acts in complex cellular systems, the dysfunction of which is only beginning to be understood<sup>101,102</sup>.

Endocytosis is a process by which cells transport molecules into the cytosol and is crucial for regulating intracellular and extracellular environments to influence cellular and circuit homeostasis. Large cytosolic protein assemblies and aged organelles undergo lysosomal degradation through macroautophagy processes that use the clathrin-mediated endocytosis of the donor membrane to engulf particles into autophagosomes<sup>102</sup>. These structures guide the degradation of proteins, and alterations in genetic risk loci associated with endocytosis (for example, the gene *BIN1*) might impair this ability<sup>8</sup>. Furthermore, synaptic activity and plasticity also rely on clathrin-mediated endocytosis at both presynaptic and postsynaptic terminals. At the presynapse, the process regulates the recycling

of neurotransmitter vesicles to maintain a pool of ready-to-release synaptic vesicles that enables fast signalling<sup>103</sup>. At the postsynapse, clathrin-mediated endocytosis mediates plasticity through receptor trafficking and signal transduction cascades<sup>36</sup>. Synaptic APP and  $\beta$ -secretase undergo internalization from the plasma membrane through clathrin-mediated endocytosis, and despite separate secretory pathways, the rapid fusion of vesicles containing APP and  $\beta$ -secretase with recycling endosomes regulates the activity-dependent cleavage of APP into amyloid- $\beta$  (ref. 29). The activity-dependent trafficking of late endosomes also mediates the proximity of  $\gamma$ -secretase to APP<sup>28</sup>. Interestingly, landmark studies in yeast that investigated amyloid- $\beta$ -induced dysfunction identified several genetic modifiers of toxicity in the endocytic pathway<sup>104</sup>. As well as genetic risk factors that can impair plasticity and enhance amyloid- $\beta$  production, amyloid- $\beta$  itself can disrupt clathrin-mediated endocytosis by altering the pattern of clathrin aggregation at the plasma membrane, which exemplifies the complexity of the feedback mechanisms that can go awry in AD<sup>104</sup>.

As with intracellular degradation and clearance, maintenance of the extracellular environment also requires the removal of debris to protect connectivity and signalling capacities. The cells that hold the main responsibility for this elimination in the brain are microglia, which survey their surroundings for unwanted and potentially disease-causing debris<sup>105</sup>. Microglia can remove such waste from the parenchyma through specialized endocytic processes called phagocytosis and pinocytosis, and the uptake of amyloid- $\beta$  by microglia seems to be an important means of amyloid- $\beta$ -aggregate degradation. Loss of the late-onset AD risk gene *TREM2* reduced phagocytic cell migration towards amyloid- $\beta$  plaques and exacerbated amyloid- $\beta$  deposition<sup>106</sup>, and phagocytic genes were downregulated in microglia following amyloid- $\beta$  accumulation<sup>107</sup>. It is therefore possible that pathogenic feed-forward processes are involved in which deficits in phagocytosis, induced either genetically or by amyloid- $\beta$ , lead to increased levels of amyloid- $\beta$  and further dysfunction. Microglia have also been shown to take up amyloid- $\beta$  oligomers through pinocytosis, a related endocytic process<sup>108</sup>. Because these oligomers induce long-term depression and have other deleterious effects on the synapse<sup>36</sup>, the loss of phagocytic functions might contribute to inefficient communication and circuit dysfunction even before the formation of overt plaques.

Emerging functions of microglia at synapses suggest that these cells also participate directly in synaptic remodelling<sup>77,105</sup>. Although considerable data implicate microglia in the development of AD, the exact role of these cells in disease progression remains uncertain. Microglia can both protect and damage circuits, which further complicates understanding of their contribution to AD<sup>77,109</sup>. For example, lipopolysaccharide-activated microglia shear inhibitory presynaptic terminals from the soma of excitatory neurons to enhance coherent firing and upregulate protective pathways<sup>109</sup>. But in response to oligomeric amyloid- $\beta$ , microglia eliminate synapses through a complement-mediated pathway in a mouse model of AD<sup>77</sup>. Although both findings indicate that microglia have a considerable role in shaping synaptic communication, it remains unclear whether they initiate neurodegeneration, confer protection or affect both outcomes. These contrasting results might reflect the distinct responses of microglia to the different phases of AD. More generally, microglia-mediated immune responses probably affect the progression of AD as well; however, these responses are also poorly understood in AD. Future studies that use omics-based screens and live-cell imaging will be necessary to disentangle the complex functions of microglia in AD.

Microglia are not the only non-neuronal cells that affect circuits: astrocytes, oligodendrocytes and endothelial cells also have considerable roles in maintaining circuit function. In AD, each of these cell types undergoes changes that probably affect neuronal activity, a topic that is thoroughly covered elsewhere<sup>9</sup>. Building on the conclusion that mapping alterations in cells can provide insights into incipient changes that underlie memory loss, it will also be crucial to understand how such changes alter circuit dynamics to affect cognition. For example, a striking finding in both amyloid- $\beta$  and tau models of AD is that the presence of aggregates increases the size of hippocampal place fields and decreases their stability<sup>110,111</sup>. The ensemble firing of place cells in normal brains generates



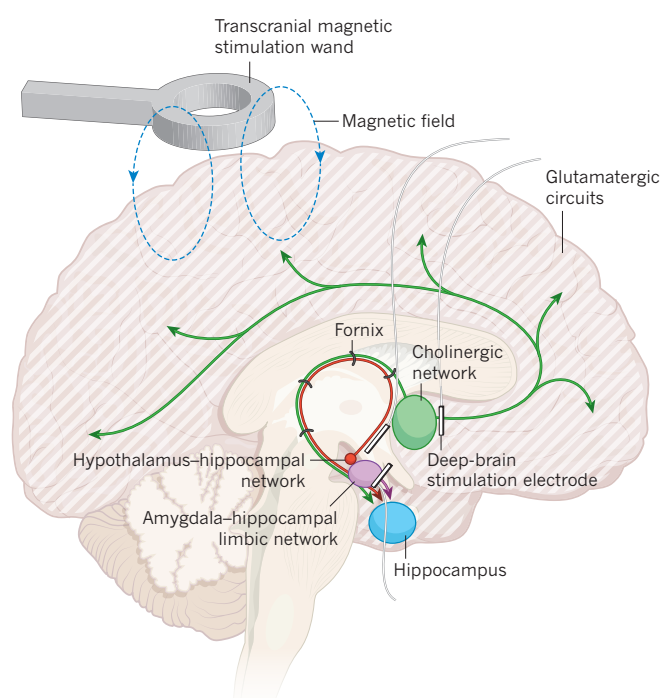
a map that promotes spatial navigation and location recognition, and alterations could explain the spatial deficits that are seen in people with AD<sup>112,113</sup>. It is unknown whether such deficits are the result of the intrinsic inability of place cells to encode information or altered network dynamics; however, reducing hyperexcitability in the dentate gyrus and CA3 subfields of the hippocampus in a mouse model of AD and in people with AD who use levetiracetam restores cognition<sup>40,41</sup>. Although the restoration was observed in an amyloid- $\beta$  mouse model of AD, it seems to be independent of any effects on amyloid- $\beta$  and correlates instead with increases in synaptic markers<sup>41</sup>. This suggests that complex alterations at both the circuit and network levels might contribute independently to cognitive decline.

### Network dysfunction predicts cognitive symptoms

The first evidence for network dysfunction in AD was the observation of cholinergic neuron loss in postmortem brains<sup>114</sup>. Cholinergic neurons are found in distinct subcortical regions of the brain and use acetylcholine as a neurotransmitter. The regions are connected to each other, and the cholinergic neurons have processes that extend as a diffuse network through the brain, which enable them to modulate the excitatory tone of circuits and to facilitate states such as sleep, wakefulness and attentiveness<sup>115</sup>. Identification of neurofibrillary-tangle formation and degeneration in cholinergic neurons<sup>116</sup> supports the cholinergic hypothesis, which states that the loss of cholinergic neurons is an event that leads to AD<sup>117</sup>. These cells affect cognitive processing through the modulation of circuit activity and coherent firing<sup>118</sup>, and drugs that target cholinergic signalling account for 75% of treatments for AD that have been approved by the US Food and Drug Administration. However, the efficacy of therapies that target cholinergic pathways is short and they might temporarily slow, but do not halt or reverse, memory decline<sup>119</sup>. Although enhancing cholinergic signalling cannot reverse the progression of AD, the limited efficacy of such treatments suggests that the loss of regulatory tone might exacerbate dysfunction and accelerate memory loss. Accumulating evidence indicates that the long-range structure of the network could make these cells particularly susceptible to deficits in axonal transport and metabolic insults that induce neurofibrillary tangles and other pathologies<sup>120</sup>. The identification of further subcortical neuromodulatory systems that show neurofibrillary tangles at the earliest stages of disease indicates that studying these networks could pinpoint specific vulnerabilities in cellular processes that change in the initial stages of AD<sup>121</sup>.

The first symptoms that people with AD report are a difficulty in remembering new information and episodic memory loss<sup>67</sup>. Neuropathological examinations show that the accumulation of neurofibrillary tangles occurs initially in the entorhinal cortex (EC) of the brain then spreads through the hippocampal formation<sup>122</sup>. Advances in PET tracers have enabled tau localization to be imaged with <sup>18</sup>F-T807; this revealed similar patterns to those determined by pathological studies of AD brains, which suggests that these tools might be useful for diagnosing AD<sup>123</sup>. In support of this, alterations in entorhinal cortex–hippocampal regions of people with neurological disorders but not AD, as well as in rats with neural lesions in the entorhinal cortex, impair learning and the retrieval of episodic memories, which shows that the anatomical findings correspond with the initial clinical symptoms of AD<sup>124,125</sup>. This indicates that processes in these regions that are disrupted in AD can alter circuits that impair learning and memory. Also, because the accumulation of neurofibrillary tangles develops subsequently in connected brain regions that control planning, emotion and navigation, the temporal sequence of increasing burden might explain the deterioration of cognitive function as AD advances<sup>126,127</sup>.

Both neurofibrillary tangles and amyloid- $\beta$  plaques propagate through the limbic system, a network of distinct brain regions that control memory and emotional behaviours<sup>122,128</sup>. The sequential involvement of limbic brain regions, and the loss of their proper function, could explain the progression of clinical symptoms, and MRI studies in people with AD demonstrate that specific atrophy in nodes of the network correlates with cognitive complaints<sup>129</sup>. Limbic regions receive regulatory inputs



**Figure 3 | Network-level treatments for AD.** Drugs that target cholinergic and glutamatergic circuits have minimal efficacy yet are the only approved treatments for AD<sup>119</sup>. Preliminary studies using deep-brain stimulation electrodes in the basal forebrain cholinergic circuit (green) and hypothalamus–hippocampal network (red) have shown promise in slowing disease progression<sup>133,144</sup>. Direct stimulation of amygdala–hippocampal circuits (purple) in rodents also improves memory, which suggests that targeting these limbic circuits directly might be another route to restoring cognition. The implementation of non-invasive technologies such as transcranial magnetic stimulation in network modification makes such treatments safer and scalable.

from subcortical structures, in particular from those that accumulate tau. Therefore, such accumulations could contribute to dysregulation of the limbic regions<sup>130</sup>. PET imaging of glucose metabolism shows a considerable reduction in glucose use in the hippocampal formation and other limbic temporal regions<sup>131</sup>. Furthermore, disruption of limbic white matter is correlated with cognitive decline in people with amnesic mild cognitive impairment<sup>132</sup>. Evidence from clinical trials supports an underlying role for limbic network dysfunction in memory loss; preliminary data from a small cohort of people with AD suggest that electrical stimulation of the fornix, a main white-matter bundle that connects limbic regions, leads to enhanced cognition<sup>133</sup>. Stimulation also induces DMN activation, which indicates that the DMN is highly interconnected with many limbic nodes<sup>133</sup>. This observation could also reconcile the temporal sequence of cognitive symptoms and point to the earliest hubs of disease generation. In the future, longitudinal, high-resolution functional MRI or PET-based investigations should examine specifically the subcortical limbic nodes to ascertain whether amyloid- $\beta$  deposits and functional disconnection can be detected there before deficits are seen in DMN function. If so, innovative methods for assessing limbic function could be useful for identifying people with the earliest stage of AD.

As well as the uncertainty about the core networks that are affected by AD, the mechanisms that underlie the spread of neurofibrillary tangles and amyloid- $\beta$  are unknown. Amyloid- $\beta$  peptides and tau proteins released at the synapse could propagate in a prion-like manner, but the biological pathways that underlie this potential movement are unclear and it remains technically difficult to demonstrate seed propagation directly. An alternative explanation for the spread of aggregates is that local circuit dysfunction causes alterations in network activity, which drives increases in amyloid- $\beta$  levels and tau aggregation through abnormal activity at downstream targets. Regional vulnerability might lead to the propagation

of disrupted cascades, which could develop into the hallmarks of AD, if left unchecked. Although increasing evidence suggests that amyloid- $\beta$  and tau spread within networks, region-specific responses to experiential factors such as ageing, diet, emotional stress and mental health might also explain the distribution of disease hallmarks<sup>11,120</sup>. The limbic system is particularly responsive to stress and to the diet, and many of its main regions, including the hippocampal formation, express high levels of receptors for stress-related glucocorticoids and gut hormones<sup>134,135</sup>. Indeed, high-fat diets and chronic stress can damage the hippocampal formation, and these environmentally responsive limbic structures might be important points at which experience, genetic risk and the onset of AD converge<sup>120,136</sup>. In a mouse model of AD, emotional stressors exacerbated the loss of dendritic spines in the hippocampal formation, and repeated exposure to a stressor in normal mice led to a neurodegenerative-like phenotype in the same brain region, as well as cognitive impairment<sup>137,138</sup>.

### Network treatments might restore memory

Therapeutic drugs that target amyloid- $\beta$  have been largely unsuccessful at restoring cognition, and most treatments for AD that have the potential to improve cognition target networks to improve or protect circuit integrity and activity<sup>40,64,119,133,139–141</sup> (Fig. 3). The main drugs prescribed to people with early stage memory loss — donepezil, rivastigmine and galantamine — promote cholinergic signalling by inhibiting acetylcholine degradation to maintain its levels in the extracellular space<sup>119</sup>, which is thought to promote neuronal activity and improve cholinergic tone. These compounds also probably affect microglia and astrocytes, both of which express acetylcholine receptors and respond to cholinergic signalling<sup>141,142</sup>. Another therapeutic drug approved for the treatment of AD is memantine, an NMDA (*N*-methyl-D-aspartate) receptor antagonist<sup>119</sup>. It is the only compound that has been approved for use in late-stage AD, and it could promote circuit connectivity by preserving synapse integrity. Memantine probably counteracts the effects of hyperactive excitatory circuits and prevents high levels of glutamate from weakening synapse strength<sup>141</sup>. Similar to molecular strategies that target amyloid- $\beta$ , these network-level treatments are mildly effective. However, they have yet to stop cognitive deterioration or to restore memory function.

Other molecular treatments to restore cellular health and to repair circuit and network functions are under development. Alongside those for amyloid- $\beta$ , antibody-based treatments that target tau aggregation and APOE have also emerged, although evaluating their efficacy using data from clinical trials is not yet possible<sup>143</sup>. These antibody-based strategies probably influence immune functions; however, the increasing evidence for non-inflammatory roles of microglia in circuit maintenance suggests that targeting other microglia processes could be a better route to mitigating cognitive impairment.

Directly targeting the activity of brain networks might also help to restore memory. Phase I clinical trials that used deep-brain stimulation techniques to directly manipulate network activity in individuals with AD reported positive memory outcomes<sup>133,144</sup>. Stimulation of the fornix also changed protein expression in animal models, which suggests that the development of non-invasive brain-stimulation strategies could be a scalable and safe route to restoring cellular health and network function<sup>145</sup>. Furthermore, using optogenetic techniques to excite cells in the hippocampus to increase the number of dendritic spines restored learning and memory<sup>139</sup>. Because similar effects were found using histone deacetylase inhibitors<sup>140</sup>, it is probable that diverse treatment strategies can leverage these effects to repair circuits. The optogenetics<sup>139</sup> and histone deacetylase<sup>140</sup> studies are particularly interesting because they used inducible models as a control for learning. They showed that, despite considerable memory impairment following neuronal degeneration, the restoration of synapses could restore recall performance. This suggests that memories in AD might be inaccessible rather than lost. Together, the efficacy of these experiments demonstrates that intervention at both the network and circuit levels can restore cellular health and circuit integrity and could provide new directions for restoring memory in AD.

### Perspective

A growing understanding of cognitive impairment in AD suggests that alterations at the genetic and cellular levels contribute to circuit dysfunction, which affects long-range network connectivity. By restoring these connections and the circuit-level and cellular-level processes that support them, it might be possible to reverse memory loss. Because no single strategy or target has been fully effective in promoting cognition, it is probable that a multitiered approach to treating AD will be necessary. Treatments that reduce amyloid- $\beta$  levels stop the cascade that triggers cellular dysfunction, and these will be especially important for people with familial AD, in which genetic mutations induce the overproduction of amyloid- $\beta$  directly. Together with a reduction in amyloid- $\beta$  to halt disease progression, restoring cognition and brain health will probably need therapeutic drugs that regulate circuit activity and stimulate neuronal communication to improve the function of long-range networks. The conventional path to these treatments requires an understanding of which disrupted processes contribute to amyloid- $\beta$  accumulation and circuit disruption. However, because circuit activity also affects cellular processes, it may be possible to circumvent this pipeline by restoring network-level and circuit-level functions in people with AD. These treatments might be able to directly improve memory function and feedback on molecular processes to re-establish cellular health, paving the way to a healthy ageing brain. ■

Received 1 May; accepted 3 August 2016.

- Hurd, M. D., Martorell, P., Delavande, A., Mullen, K. J. & Langa, K. M. Monetary costs of dementia in the United States. *N. Engl. J. Med.* **368**, 1326–1334 (2013).
- Hyman, B. T. et al. National Institute on Aging—Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement.* **8**, 1–13 (2012).
- Goate, A. et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* **349**, 704–706 (1991). **One of the first demonstrations that mutations in the gene *APP* correlate with familial AD, suggesting a prominent role for amyloid- $\beta$  processing in the aetiology of AD.**
- Schellenberg, G. et al. Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. *Science* **258**, 668–671 (1992).
- Levy-Lahad, E. et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* **269**, 973–977 (1995).
- Hardy, J. & Higgins, G. Alzheimer's disease: the amyloid cascade hypothesis. *Science* **256**, 184–185 (1992). **First article to synthesize the available data to formally propose that amyloid- $\beta$  leads to varied neuronal disruption and cognitive impairment in AD.**
- Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).
- Lambert, J. C. et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genet.* **45**, 1452–1458 (2013). By aggregating data, this paper confirmed previous genetic risk factors associated with late-onset AD and also identified new loci that might increase susceptibility to the disease.
- De Strooper, B. & Karran, E. The cellular phase of Alzheimer's disease. *Cell* **164**, 603–615 (2016).
- Palop, J. J. & Mucke, L. Amyloid- $\beta$ -induced neuronal dysfunction in Alzheimer's disease: from synapses toward neural networks. *Nature Neurosci.* **13**, 812–818 (2010).
- Reitz, C. & Mayeux, R. Alzheimer disease: epidemiology, diagnostic criteria, risk factors and biomarkers. *Biochem. Pharmacol.* **88**, 640–651 (2014).
- Hardy, J. The Alzheimer family of diseases: many etiologies, one pathogenesis? *Proc. Natl Acad. Sci. USA* **94**, 2095–2097 (1997).
- Rovelet-Lecrux, A. et al. APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nature Genet.* **38**, 24–26 (2006).
- Glenner, G. G. & Wong, C. W. Alzheimer's disease and Down's syndrome: sharing of a unique cerebrovascular amyloid fibril protein. *Biochem. Biophys. Res. Commun.* **122**, 1131–1135 (1984).
- Hutton, M. et al. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* **393**, 702–705 (1998).
- Jonsson, T. et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* **488**, 96–99 (2012).
- Corder, E. H. et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **261**, 921–923 (1993).
- Morris, J. C. et al. APOE predicts amyloid- $\beta$  but not tau Alzheimer pathology in cognitively normal aging. *Ann. Neurol.* **67**, 122–131 (2010).
- Verghese, P. B. et al. ApoE influences amyloid- $\beta$  (A $\beta$ ) clearance despite minimal apoE/A $\beta$  association in physiological conditions. *Proc. Natl Acad. Sci. USA* **110**, E1807–E1816 (2013).
- Kayed, R. & Lasagna-Reeves, C. Molecular mechanisms of amyloid oligomers toxicity. *J. Alzheimers Dis.* **33**, S67–S78 (2013).
- Yankner, B. A., Duffy, L. K. & Kirschner, D. A. Neurotrophic and neurotoxic effects of amyloid beta protein: reversal by tachykinin neuropeptides. *Science* **250**, 279–282 (1990).
- Jack, C. R. et al. Serial PIB and MRI in normal, mild cognitive impairment



- and Alzheimer's disease: implications for sequence of pathological events in Alzheimer's disease. *Brain* **132**, 1355–1365 (2009).
- Showed how certain events unfold in people with AD and contributed to an understanding of the temporal disconnection between amyloid- $\beta$  deposition and cognitive impairment.**
23. Shankar, G. M. *et al.* Natural oligomers of the Alzheimer amyloid- $\beta$  protein induce reversible synapse loss by modulating an NMDA-type glutamate receptor-dependent signaling pathway. *J. Neurosci.* **27**, 2866–2875 (2007).
  24. Šišková, Z. *et al.* Dendritic structural degeneration is functionally linked to cellular hyperexcitability in a mouse model of Alzheimer's disease. *Neuron* **84**, 1023–1033 (2014).
  25. Palop, J. J. *et al.* Aberrant excitatory neuronal activity and compensatory remodeling of inhibitory hippocampal circuits in mouse models of Alzheimer's disease. *Neuron* **55**, 697–711 (2007).
  - Demonstrated that amyloid- $\beta$  and AD-like features could induce hyperactivity in regions of the brain, challenging the idea that neurodegeneration leads to reduced neuronal activity and highlighting the complexity of changes seen in AD.**
  26. Cirrito, J. R. *et al.* Synaptic activity regulates interstitial fluid amyloid- $\beta$  levels *in vivo*. *Neuron* **48**, 913–922 (2005).
  - First to indicate a considerable physiological role for amyloid- $\beta$  in the brain, suggesting it could have important functions that influence AD phenotypes.**
  27. Wei, W. *et al.* Amyloid  $\beta$  from axons and dendrites reduces local spine number and plasticity. *Nature Neurosci.* **13**, 190–196 (2010).
  28. Wu, J. *et al.* Arc/Arg3.1 regulates an endosomal pathway essential for activity-dependent  $\beta$ -amyloid generation. *Cell* **147**, 615–628 (2011).
  29. Das, U. *et al.* Activity-induced convergence of APP and BACE-1 in acidic microdomains via an endocytosis-dependent pathway. *Neuron* **79**, 447–460 (2013).
  30. Dougherty, J. J., Wu, J. & Nichols, R. A.  $\beta$ -Amyloid regulation of presynaptic nicotinic receptors in rat hippocampus and neocortex. *J. Neurosci.* **23**, 6740–6747 (2003).
  31. Abramov, E. *et al.* Amyloid- $\beta$  as a positive endogenous regulator of release probability at hippocampal synapses. *Nature Neurosci.* **12**, 1567–1576 (2009).
  32. Snyder, E. M. *et al.* Regulation of NMDA receptor trafficking by amyloid- $\beta$ . *Nature Neurosci.* **8**, 1051–1058 (2005).
  33. Roselli, F. *et al.* Soluble  $\beta$ -amyloid<sub>1–40</sub> induces NMDA-dependent degradation of postsynaptic density-95 at glutamatergic synapses. *J. Neurosci.* **25**, 11061–11070 (2005).
  34. Walsh, D. M. *et al.* Naturally secreted oligomers of amyloid  $\beta$  protein potently inhibit hippocampal long-term potentiation *in vivo*. *Nature* **416**, 535–539 (2002).
  35. Li, S. *et al.* Soluble oligomers of amyloid  $\beta$  protein facilitate hippocampal long-term depression by disrupting neuronal glutamate uptake. *Neuron* **62**, 788–801 (2009).
  36. Hsieh, H. *et al.* AMPAR removal underlies  $\beta$ -induced synaptic depression and dendritic spine loss. *Neuron* **52**, 831–843 (2006).
  37. Verret, E. *et al.* Inhibitory interneuron deficit links altered network activity and cognitive dysfunction in Alzheimer model. *Cell* **149**, 708–721 (2012).
  38. Vossel, K. A. *et al.* Seizures and epileptiform activity in the early stages of Alzheimer disease. *JAMA Neurol.* **70**, 1158–1166 (2013).
  - Provided an early demonstration of neuronal hyperactivity in people with AD and confirmed findings made originally in rodent models of AD, suggesting that such models can recapitulate facets of AD accurately, despite limitations.**
  39. Volicer, L., Smith, S. & Volicer, B. J. Effect of seizures on progression of dementia of the Alzheimer type. *Dementia* **6**, 258–263 (1995).
  40. Bakker, A., Albert, M. S., Krauss, G., Speck, C. L. & Gallagher, M. Response of the medial temporal lobe network in amnesic mild cognitive impairment to therapeutic intervention assessed by fMRI and memory task performance. *Neuroimage Clin.* **7**, 688–698 (2015).
  41. Sanchez, P. E. *et al.* Levetiracetam suppresses neuronal network dysfunction and reverses synaptic and cognitive deficits in an Alzheimer's disease model. *Proc. Natl Acad. Sci. USA* **109**, E2895–E2903 (2012).
  42. Busche, M. A. *et al.* Rescue of long-range circuit dysfunction in Alzheimer's disease models. *Nature Neurosci.* **18**, 1623–1630 (2015).
  43. Thal, D. R., Rüb, U., Orantes, M. & Braak, H. Phases of  $\beta$ -deposition in the human brain and its relevance for the development of AD. *Neurology* **58**, 1791–1800 (2002).
  44. Klunk, W. E. *et al.* Imaging brain amyloid in Alzheimer's disease with Pittsburgh compound-B. *Ann. Neurol.* **55**, 306–319 (2004).
  - This important paper was the first to describe the PET imaging of amyloids in people with AD.**
  45. Wong, D. F. *et al.* *In vivo* imaging of amyloid deposition in Alzheimer disease using the radioligand <sup>18</sup>F-AV-45 (florbetapir F 18). *J. Nucl. Med.* **51**, 913–920 (2010).
  46. Jack, C. R. *et al.* Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer's disease. *Brain* **133**, 3336–3348 (2010).
  47. Forsberg, A. *et al.* PET imaging of amyloid deposition in patients with mild cognitive impairment. *Neurobiol. Aging* **29**, 1456–1465 (2008).
  48. Albert, M. S. *et al.* The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* **7**, 270–279 (2011).
  49. Greicius, M. D., Srivastava, G., Reiss, A. L. & Menon, V. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc. Natl Acad. Sci. USA* **101**, 4637–4642 (2004).
  50. Buckner, R. L. Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *J. Neurosci.* **25**, 7709–7717 (2005).
  51. Sperling, R. A., LaViolette, P. & O'Keefe, K. Amyloid deposition is associated with impaired default network function in older persons without dementia. *Neuron* **63**, 178–188 (2009).
  52. Mormino, E. C. *et al.* Synergistic effect of  $\beta$ -amyloid and neurodegeneration on cognitive decline in clinically normal individuals. *JAMA Neurol.* **71**, 1379–1385 (2014).
  53. Villemagne, V. L. *et al.* Longitudinal assessment of A $\beta$  and cognition in aging and Alzheimer disease. *Ann. Neurol.* **69**, 181–192 (2011).
  54. Damoiseaux, J. S., Prater, K. E., Miller, B. L. & Greicius, M. D. Functional connectivity tracks clinical deterioration in Alzheimer's disease. *Neurobiol. Aging* **33**, e19–e28.e30 (2012).
  55. Harris, J. A. *et al.* Transsynaptic progression of amyloid- $\beta$ -induced neuronal dysfunction within the entorhinal–hippocampal network. *Neuron* **68**, 428–441 (2010).
  56. Jucker, M. & Walker, L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45–51 (2013).
  57. Fleisher, A. S. *et al.* Phase 2 safety trial targeting amyloid  $\beta$  production with a  $\gamma$ -secretase inhibitor in Alzheimer disease. *Arch. Neurol.* **65**, 1031–1038 (2008).
  58. May, P. C. *et al.* The Potent BACE1 inhibitor LY2886721 elicits robust central A $\beta$  pharmacodynamic responses in mice, dogs, and humans. *J. Neurosci.* **35**, 1199–1210 (2015).
  59. De Strooper, B. Lessons from a failed  $\gamma$ -secretase Alzheimer trial. *Cell* **159**, 721–726 (2014).
  60. Vassar, R. BACE1 inhibitor drugs in clinical trials for Alzheimer's disease. *Alzheimers Res. Ther.* **6**, 89 (2014).
  61. Doody, R. S. *et al.* A phase 3 trial of semagacestat for treatment of Alzheimer's disease. *N. Engl. J. Med.* **369**, 341–350 (2013).
  62. Orgogozo, J. M. *et al.* Subacute meningoencephalitis in a subset of patients with AD after A $\beta$ <sub>42</sub> immunization. *Neurology* **61**, 46–54 (2003).
  63. Farlow, M. R. *et al.* Long-term treatment with active A $\beta$  immunotherapy with CAD106 in mild Alzheimer's disease. *Alzheimers Res. Ther.* **7**, 23 (2015).
  64. Holmes, C. *et al.* Long-term effects of A $\beta$ <sub>42</sub> immunisation in Alzheimer's disease: follow-up of a randomised, placebo-controlled phase I trial. *Lancet* **372**, 216–223 (2008).
  65. Siemers, E. R. *et al.* Phase 3 solanezumab trials: secondary outcomes in mild Alzheimer's disease patients. *Alzheimers Dement.* **12**, 110–120 (2016).
  66. J. Sevigny *et al.* The antibody aducanumab reduces A $\beta$  plaques in Alzheimer's disease. *Nature* **537**, 50–56 (2016).
  67. McKhann, G. M. *et al.* The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* **7**, 263–269 (2011).
  68. van der Kant, R. & Goldstein, L. S. B. Cellular functions of the amyloid precursor protein from development to dementia. *Dev. Cell* **32**, 502–515 (2015).
  69. Duggan, S. P. & McCarthy, J. V. Beyond  $\gamma$ -secretase activity: the multifunctional nature of presenilins in cell signalling pathways. *Cell. Signal.* **28**, 1–11 (2016).
  70. Müller, U. *et al.* Behavioral and anatomical deficits in mice homozygous for a modified  $\beta$ -amyloid precursor protein gene. *Cell* **79**, 755–765 (1994).
  71. Dawson, G. R. *et al.* Age-related cognitive deficits, impaired long-term potentiation and reduction in synaptic marker density in mice lacking the  $\beta$ -amyloid precursor protein. *Neuroscience* **90**, 1–13 (1999).
  72. Xia, D. *et al.* Presenilin-1 knockin mice reveal loss-of-function mechanism for familial Alzheimer's disease. *Neuron* **85**, 967–981 (2015).
  73. Nelson, O. *et al.* Familial Alzheimer disease-linked mutations specifically disrupt Ca<sup>2+</sup> leak function of presenilin 1. *J. Clin. Invest.* **117**, 1230–1239 (2007).
  74. Fabelo, N. *et al.* Altered lipid composition in cortical lipid rafts occurs at early stages of sporadic Alzheimer's disease and facilitates APP/BACE1 interactions. *Neurobiol. Aging* **35**, 1801–1812 (2014).
  75. Sakae, N. *et al.* ABCA7 deficiency accelerates amyloid- $\beta$  generation and Alzheimer's neuronal pathology. *J. Neurosci.* **36**, 3848–3859 (2016).
  76. Miller, S. E. *et al.* CALM regulates clathrin-coated vesicle size and maturation by directly sensing and driving membrane curvature. *Dev. Cell* **33**, 163–175 (2015).
  77. Hong, S. *et al.* Complement and microglia mediate early synapse loss in Alzheimer mouse models. *Science* **352**, 712–716 (2016).
  78. De Jager, P. L. *et al.* Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nature Neurosci.* **17**, 1156–1163 (2014).
  79. Yu, L. *et al.* Association of brain DNA methylation in SORL1, ABCA7, HLA-DRB5, SLC24A4, and BIN1 with pathological diagnosis of Alzheimer disease. *JAMA Neurol.* **72**, 15–24 (2015).
  80. Gijóneska, E. *et al.* Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* **518**, 365–369 (2015).
  81. Gräff, J. *et al.* An epigenetic blockade of cognitive functions in the neurodegenerating brain. *Nature* **483**, 222–226 (2012).
  82. Lu, T. *et al.* REST and stress resistance in ageing and Alzheimer's disease. *Nature* **507**, 448–454 (2014).
  83. Jo, D.-G. *et al.* Evidence that  $\gamma$ -secretase mediates oxidative stress-induced  $\beta$ -secretase expression in Alzheimer's disease. *Neurobiol. Aging* **31**, 917–925 (2010).
  84. Mecocci, P., MacGarvey, U. & Beal, M. F. Oxidative damage to mitochondrial DNA is increased in Alzheimer's disease. *Ann. Neurol.* **36**, 747–751 (1994).
  85. Frost, B., Hemberg, M., Lewis, J. & Feany, M. B. Tau promotes neurodegeneration through global chromatin relaxation. *Nature Neurosci.* **17**, 357–366 (2014).
  86. Kim, D. *et al.* Deregulation of HDAC1 by p25/Cdk5 in neurotoxicity. *Neuron* **60**, 803–817 (2008).
  87. Anderson, A. J., Stoltzner, S., Lai, F., Su, J. & Nixon, R. A. Morphological and biochemical assessment of DNA damage and apoptosis in Down syndrome and Alzheimer disease, and effect of postmortem tissue archival on TUNEL. *Neurobiol. Aging* **21**, 511–524 (2000).
  88. Suberbielle, E. *et al.* Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid- $\beta$ . *Nature Neurosci.* **16**, 613–621 (2013).
  89. Lu, T. *et al.* Gene regulation and DNA damage in the ageing human brain. *Nature* **429**, 883–891 (2004).

90. Madabhushi, R. *et al.* Activity-induced DNA breaks govern the expression of neuronal early-response genes. *Cell* **161**, 1592–1605 (2015).
91. LaFerla, F. M. Calcium dyshomeostasis and intracellular signalling in Alzheimer's disease. *Nature Rev. Neurosci.* **3**, 862–872 (2002).
92. Seo, J. *et al.* Activity-dependent p25 generation regulates synaptic plasticity and A $\beta$ -induced cognitive impairment. *Cell* **157**, 486–498 (2014).
93. Lindwall, G. & Cole, R. D. Phosphorylation affects the ability of tau protein to promote microtubule assembly. *J. Biol. Chem.* **259**, 5301–5305 (1984).
94. Bramblett, G. T. *et al.* Abnormal tau phosphorylation at Ser396 in Alzheimer's disease recapitulates development and contributes to reduced microtubule binding. *Neuron* **10**, 1089–1099 (1993).
95. Sontag, E., Nunbhakdi-Craig, V., Lee, G., Bloom, G. S. & Mumby, M. C. Regulation of the phosphorylation state and microtubule-binding activity of tau by protein phosphatase 2A. *Neuron* **17**, 1201–1207 (1996).
96. Vogelsberg-Ragaglia, V., Schuck, T., Trojanowski, J. Q. & Lee, V. M. PP2A mRNA expression is quantitatively decreased in Alzheimer's disease hippocampus. *Exp. Neurol.* **168**, 402–412 (2001).
97. Cruz, J. C., Tseng, H. C., Goldman, J. A., Shih, H. & Tsai, L. H. Aberrant Cdk5 activation by p25 triggers pathological events leading to neurodegeneration and neurofibrillary tangles. *Neuron* **40**, 471–483 (2003).
98. Yamada, K. *et al.* Neuronal activity regulates extracellular tau *in vivo*. *J. Exp. Med.* **211**, 387–393 (2014).
99. Pooler, A. M., Phillips, E. C., Lau, D. H. W., Noble, W. & Hanger, D. P. Physiological release of endogenous tau is stimulated by neuronal activity. *EMBO Rep.* **14**, 389–394 (2013).
100. Wu, J. W. *et al.* Neuronal activity enhances tau propagation and tau pathology *in vivo*. *Nature Neurosci.* **19**, 1085–1092 (2016).
101. Cohen, T. J. *et al.* The acetylation of tau inhibits its function and promotes pathological tau aggregation. *Nature Commun.* **2**, 252 (2011).
102. Moreau, K. *et al.* PICALM modulates autophagy activity and tau accumulation. *Nature Commun.* **5**, 4998 (2014).
103. Kononenko, N. L. & Haucke, V. Molecular mechanisms of presynaptic membrane retrieval and synaptic vesicle reformation. *Neuron* **85**, 484–496 (2015).
104. Treusch, S. *et al.* Functional links between A $\beta$  toxicity, endocytic trafficking, and Alzheimer's disease risk factors in yeast. *Science* **334**, 1241–1245 (2011).
105. Crotti, A. & Ransohoff, R. M. Microglial physiology and pathophysiology: insights from genome-wide transcriptional profiling. *Immunity* **44**, 505–515 (2016).
106. Wang, Y. *et al.* TREM2 lipid sensing sustains the microglial response in an Alzheimer's disease model. *Cell* **160**, 1061–1071 (2015).
- The first demonstration that AD genetic risk loci affect microglia, which play an important part in modulating the AD state.**
107. Orre, M. *et al.* Isolation of glia from Alzheimer's mice reveals inflammation and dysfunction. *Neurobiol. Aging* **35**, 2746–2760 (2014).
108. Mandrekar, S. *et al.* Microglia mediate the clearance of soluble A $\beta$  through fluid phase macropinocytosis. *J. Neurosci.* **29**, 4252–4262 (2009).
109. Chen, Z. *et al.* Microglial displacement of inhibitory synapses provides neuroprotection in the adult brain. *Nature Commun.* **5**, 4486 (2014).
110. Cacucci, F., Yi, M., Willis, T. J., Chapman, P. & O'Keefe, J. Place cell firing correlates with memory deficits and amyloid plaque burden in Tg2576 Alzheimer mouse model. *Proc. Natl Acad. Sci. USA* **105**, 7863–7868 (2008).
111. Ciupek, S. M., Cheng, J., Ali, Y. O., Lu, H.-C. & Ji, D. Progressive functional impairments of hippocampal neurons in a tauopathy mouse model. *J. Neurosci.* **35**, 8118–8131 (2015).
112. Morris, R. G., Garrud, P., Rawlins, J. N. & O'Keefe, J. Place navigation impaired in rats with hippocampal lesions. *Nature* **297**, 681–683 (1982).
113. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space. *Science* **261**, 1055–1058 (1993).
114. Whitehouse, P. J., Price, D. L., Clark, A. W., Coyle, J. T. & DeLong, M. R. Alzheimer disease: evidence for selective loss of cholinergic neurons in the nucleus basalis. *Ann. Neurol.* **10**, 122–126 (1981).
115. Metherate, R., Cox, C. L. & Ashe, J. H. Cellular bases of neocortical activation: modulation of neural oscillations by the nucleus basalis and endogenous acetylcholine. *J. Neurosci.* **12**, 4701–4711 (1992).
116. Mesulam, M., Shaw, P., Mash, D. & Weintraub, S. Cholinergic nucleus basalis tauopathy emerges early in the aging-MCI-AD continuum. *Ann. Neurol.* **55**, 815–828 (2004).
117. Bartus, R. T., Dean, R. L., Beer, B. & Lippa, A. The cholinergic hypothesis of geriatric memory dysfunction. *Science* **217**, 408–414 (1982).
118. Hangya, B., Ranade, S. P., Lorenc, M. & Kepecs, A. Central cholinergic neurons are rapidly recruited by reinforcement feedback. *Cell* **162**, 1155–1168 (2015).
119. Zhu, C. W. *et al.* Long-term associations between cholinesterase inhibitors and memantine use and health outcomes among patients with Alzheimer's disease. *Alzheimers Dement.* **9**, 733–740 (2013).
120. Mattson, M. P. Pathways towards and away from Alzheimer's disease. *Nature* **430**, 631–639 (2004).
121. Braak, H. & Del Tredici, K. The pathological process underlying Alzheimer's disease in individuals under thirty. *Acta Neuropathol.* **121**, 171–181 (2011).
122. Braak, H. & Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* **82**, 239–259 (1991).
- The first comprehensive hierarchical staging of AD, which suggested a stereotypical path of progression and highlighted correlations between pathological observations of the brain and cognitive decline.**
123. Johnson, K. A. *et al.* Tau positron emission tomographic imaging in aging and early Alzheimer disease. *Ann. Neurol.* **79**, 110–119 (2016).
124. Steffenach, H. A., Witter, M., Moser, M. B. & Moser, E. I. Spatial memory in the rat requires the dorsolateral band of the entorhinal cortex. *Neuron* **45**, 301–313 (2005).
125. Vargha-Khadem, F. *et al.* Differential effects of early hippocampal pathology on episodic and semantic memory. *Science* **277**, 376–380 (1997).
126. Schöll, M. *et al.* PET imaging of tau deposition in the aging human brain. *Neuron* **89**, 971–982 (2016).
127. Ossenkoppele, R. *et al.* Tau PET patterns mirror clinical and neuroanatomical variability in Alzheimer's disease. *Brain* **139**, 1551–1567 (2016).
128. Ye, L. *et al.* Progression of seed-induced A $\beta$  deposition within the limbic connectome. *Brain Pathol.* **25**, 743–752 (2015).
129. Bruen, P. D., McGeown, W. J., Shanks, M. F. & Venneri, A. Neuroanatomical correlates of neuropsychiatric symptoms in Alzheimer's disease. *Brain* **131**, 2455–2463 (2008).
130. Iba, M. *et al.* Tau pathology spread in PS19 tau transgenic mice following locus coeruleus (LC) injections of synthetic tau fibrils is determined by the LC's afferent and efferent connections. *Acta Neuropathol.* **130**, 349–362 (2015).
131. Nestor, P. J., Fryer, T. D., Smielewski, P. & Hodges, J. R. Limbic hypometabolism in Alzheimer's disease and mild cognitive impairment. *Ann. Neurol.* **54**, 343–351 (2003).
132. Fletcher, E., Carmichael, O., Pasternak, O., Maier-Hein, K. H. & DeCarli, C. Early brain loss in circuits affected by Alzheimer's disease is predicted by fornix microstructure but may be independent of gray matter. *Front. Aging Neurosci.* **6**, 106 (2014).
133. Laxton, A. W. *et al.* A phase I trial of deep brain stimulation of memory circuits in Alzheimer's disease. *Ann. Neurol.* **68**, 521–534 (2010).
- This paper describes the feasibility of deep-brain stimulation in people and targets regions that are not typically described in studies of disrupted processes in AD; it suggests that complex brain-wide network effects occur in AD.**
134. McEwen, B. S. Central effects of stress hormones in health and disease: understanding the protective and damaging effects of stress and stress mediators. *Eur. J. Pharmacol.* **583**, 174–185 (2008).
135. Diano, S. *et al.* Ghrelin controls hippocampal spine synapse density and memory performance. *Nature Neurosci.* **9**, 381–388 (2006).
136. Lupien, S. J. *et al.* Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature Neurosci.* **1**, 69–73 (1998).
137. Baglietto-Vargas, D. *et al.* Short-term modern life-like stress exacerbates A $\beta$ -pathology and synapse loss in 3xTg-AD mice. *J. Neurochem.* **134**, 915–926 (2015).
138. Rei, D. *et al.* Basolateral amygdala bidirectionally modulates stress-induced hippocampal learning and memory deficits through a p25/Cdk5-dependent pathway. *Proc. Natl Acad. Sci. USA* **112**, 7291–7296 (2015).
139. Roy, D. S. *et al.* Memory retrieval by activating engram cells in mouse models of early Alzheimer's disease. *Nature* **531**, 508–512 (2016).
140. Fischer, A., Sananbenesi, F., Wang, X., Dobbin, M. & Tsai, L.-H. Recovery of learning and memory is associated with chromatin remodelling. *Nature* **447**, 178–182 (2007).
- One of the first papers to highlight the importance of epigenetic alterations in AD and the first to demonstrate experimentally that memories can be recovered through treatment, even after considerable neurodegeneration has occurred.**
141. Talantova, M. *et al.* A $\beta$  induces astrocytic glutamate release, extrasynaptic NMDA receptor activation, and synaptic loss. *Proc. Natl Acad. Sci. USA* **110**, E2518–E2527 (2013).
142. Shytle, R. D. *et al.* Cholinergic modulation of microglial activation by  $\alpha 7$  nicotinic receptors. *J. Neurochem.* **89**, 337–343 (2004).
143. Pedersen, J. T. & Sigurdsson, E. M. Tau immunotherapy for Alzheimer's disease. *Trends Mol. Med.* **21**, 394–402 (2015).
144. Kuhn, J. *et al.* Deep brain stimulation of the nucleus basalis of meynert in early stage of Alzheimer's dementia. *Brain Stimul.* **8**, 838–839 (2015).
145. Sankar, T. *et al.* Deep brain stimulation influences brain structure in Alzheimer's disease. *Brain Stimul.* **8**, 645–654 (2015).
146. Hamelin, L. *et al.* Early and protective microglial activation in Alzheimer's disease: a prospective study using  $^{18}\text{F}$ -DPA-714 PET imaging. *Brain* **139**, 1252–1264 (2016).
147. Nhan, H. S., Chiang, K. & Koo, E. H. The multifaceted nature of amyloid precursor protein and its proteolytic fragments: friends and foes. *Acta Neuropathol.* **129**, 1–19 (2015).
148. Lammich, S. *et al.* Constitutive and regulated  $\alpha$ -secretase cleavage of Alzheimer's amyloid precursor protein by a disintegrin metalloprotease. *Proc. Natl Acad. Sci. USA* **96**, 3922–3927 (1999).
149. De Jonghe, C. *et al.* Pathogenic APP mutations near the  $\gamma$ -secretase cleavage site differentially affect A $\beta$  secretion and APP C-terminal fragment stability. *Hum. Mol. Genet.* **10**, 1665–1671 (2001).
150. Willem, M. *et al.*  $\eta$ -Secretase processing of APP inhibits neuronal activity in the hippocampus. *Nature* **526**, 443–447 (2015).

**Acknowledgements** We thank the US National Institutes of Health for grants R01 NS051874, R01 NS078839, RF1 AG042978 and RF1 AG047661 in support of L.-H.T. We thank the Barbara J. Weedon Fellowship and Norman B. Leventhal Fellowship for supporting R.G.C. and the Human Frontier Science Program for supporting J.P. We also thank the JPB Foundation, the Belfer Neurodegeneration Consortium, the Glenn Foundation for Medical Research, the Cure Alzheimer's Fund and the Alana Foundation for support of L.-H.T. and for continued championship of ageing and neurodegenerative disease research. We thank C. Yao for contributions to the original figure artwork. Last, we express profound gratitude to A. Watson, H. Meharena, W. Raja and N. Dedic for insightful comments on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this paper at [go.nature.com/2exkigg](http://go.nature.com/2exkigg). Correspondence should be addressed to L.-H.T. (lhtsai@mit.edu).

**Reviewer Information** *Nature* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.



# Decoding ALS: from genes to mechanism

J. Paul Taylor<sup>1</sup>, Robert H. Brown Jr<sup>3</sup> & Don W. Cleveland<sup>4,5</sup>

**Amyotrophic lateral sclerosis (ALS) is a progressive and uniformly fatal neurodegenerative disease. A plethora of genetic factors have been identified that drive the degeneration of motor neurons in ALS, increase susceptibility to the disease or influence the rate of its progression. Emerging themes include dysfunction in RNA metabolism and protein homeostasis, with specific defects in nucleocytoplasmic trafficking, the induction of stress at the endoplasmic reticulum and impaired dynamics of ribonucleoprotein bodies such as RNA granules that assemble through liquid–liquid phase separation. Extraordinary progress in understanding the biology of ALS provides new reasons for optimism that meaningful therapies will be identified.**

Amyotrophic lateral sclerosis (ALS) is a prototypical neurodegenerative disease that is characterized by the progressive degeneration of motor neurons in the brain and spinal cord. The condition was first described by the neurologist Jean-Martin Charcot, and its name reflects both the degeneration of corticospinal motor neurons, the descending axons of which in the lateral spinal cord seem scarred (lateral sclerosis), and the demise of spinal motor neurons, with secondary denervation and muscle wasting (amyotrophy). Corticospinal neurons make direct or indirect connections with spinal motor neurons, which innervate skeletal muscles and trigger their contraction (Fig. 1a). This Review summarizes the clinical and pathological features of ALS and describes how discoveries in ALS genetics have illuminated important themes in the molecular pathophysiology of the disease.

ALS is known as Lou Gehrig's disease in the United States and as motor neuron disease in the United Kingdom. Although onset of the disease occurs commonly in mid-adulthood (at a mean age of 55 years), ALS might begin as early as in the first or second decade of life or could even emerge in later life. Similar to most neurodegenerative diseases, it starts focally and spreads: symptoms that start as subtle cramping or weakness in the limbs or bulbar muscles progress to the paralysis of almost all skeletal muscles. Some subsets of motor neurons, including those that innervate the extraocular muscles or sphincters, are spared until late in the progression of the disease. However, ALS is invariably fatal. Death occurs typically 3–5 years after diagnosis, although some forms of the disease demonstrate protracted survival.

ALS is an orphan disease that is diagnosed in 1–2 individuals per 100,000 each year in most countries; the prevalence of ALS is about 5 cases per 100,000 people, which reflects the rapid lethality of the disease<sup>1</sup>. In the United States and United Kingdom, ALS causes more than 1 in 500 deaths in adults, a statistic that suggests that more than 15 million people who are alive at present will succumb to the disease.

## The clinical manifestations of ALS

Considerable heterogeneity exists in the general rubric of ALS. Clinical subsets of the disease are distinguished by the involvement of different sets of motor neurons or different regions of the body. Depending on the location of the main pathology, those affected might develop weakness with flaccidity and atrophy of the limbs (known as progressive muscular atrophy, which mainly affects spinal neurons or lower motor neurons),

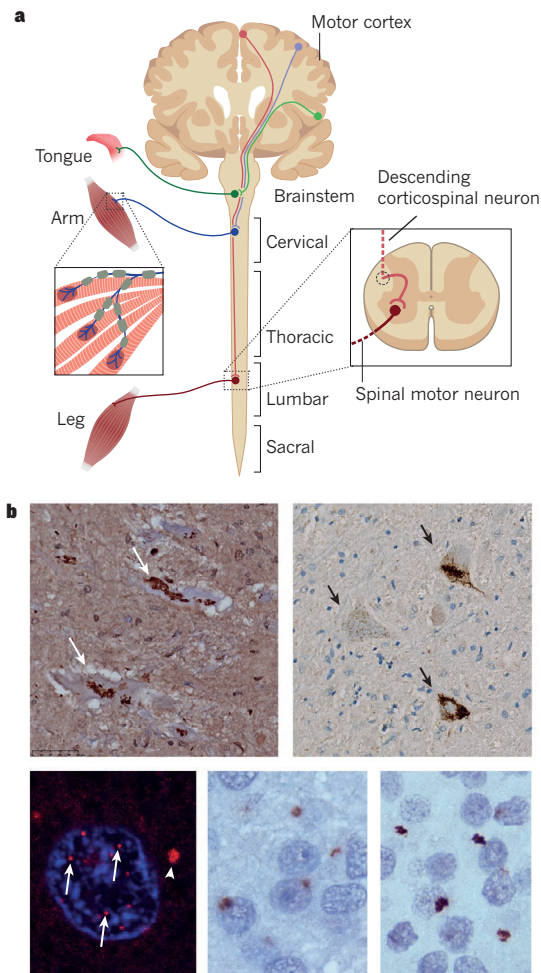
prominent hyperreflexia and spasticity with increased limb tone but little muscle atrophy (known as primary lateral sclerosis, which affects corticospinal motor neurons with limited involvement of spinal motor neurons), tongue atrophy with thickness of speech and difficulty swallowing (known as bulbar ALS, which affects brainstem motor neurons that serve the muscles of tongue movement, chewing, swallowing and articulation) or slow and highly dysfunctional speech and swallowing in the absence of tongue atrophy, often accompanied by the accentuation of emotional reflexes (known as pseudobulbar palsy, which affects cortical frontobulbar motor neurons). Importantly, ALS shares clinical and pathological features with several other adult-onset degenerative disorders, including, most frequently, frontotemporal dementia (FTD), which could constitute a clinical spectrum (Box 1).

## Genetic contributions to ALS

About 10% of ALS cases are transmitted in families, almost always as a dominant trait and frequently with high penetrance. The first genetic mutations found to cause ALS, reported in 1993, affected the gene *SOD1* (ref. 2) and more than 50 further potential ALS genes have been published since, although validating the causality of specific variants remains a challenge. By applying rigorous criteria, a list of genes with mutations that are implicated unequivocally in the pathogenesis of ALS can be generated (Table 1). These genes can be grouped into several loose categories: genes that alter proteostasis and protein quality control; genes that perturb aspects of RNA stability, function and metabolism; and genes that disturb cytoskeletal dynamics in the motor neuron axon and distal terminal. The mutations involved are mostly missense substitutions, although the genetic lesion in *C9orf72* is an enormous expansion of an intronic hexanucleotide repeat.

Although sporadic ALS should refer strictly to disease that presents without a family history of ALS, this term is sometimes mistakenly used to refer to ALS that occurs without a genetic basis. Technological advances that facilitate broad DNA sequencing in people with sporadic ALS have revealed that genetic variants in established ALS genes are not infrequent. For example, it is now evident that 1–3% of sporadic cases of ALS are caused by missense mutations in *SOD1* (ref. 3) and another 5% or more are caused by intronic expansions in *C9orf72* (ref. 4). Pathogenic mutations in other ALS genes, including *TARDBP*, *FUS*, *HNRNPA1*, *SQSTM1*, *VCP*, *OPTN* and *PFN1*, have also been identified

<sup>1</sup>Howard Hughes Medical Institute and the Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. <sup>2</sup>Department of Neurology, University of Massachusetts Medical School, Worcester, Massachusetts 01655, USA. <sup>3</sup>Ludwig Institute for Cancer Research, University of California, San Diego, La Jolla, California 92093, USA. <sup>4</sup>Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California 92093, USA.



**Figure 1 | Components of the nervous system that are affected by ALS.** **a**, ALS mainly affects the descending corticospinal motor neurons (upper motor neurons) that project from the motor cortex into synapses in the brainstem and spinal cord, and the bulbar or spinal motor neurons (lower motor neurons) that project into skeletal muscles. **b**, Subtypes of ALS show typical pathological features: SOD1 aggregates (arrows) in spinal motor neurons in SOD1-related familial ALS (top left); TDP-43 redistribution to cytoplasmic inclusions (arrows) in spinal motor neurons in sporadic ALS (top right); RNA foci in the nucleus (arrows) and the cytoplasm (arrowhead) of a cortical neuron affected by C9 ALS-FTD (bottom left); GA (bottom centre) and GR (bottom right) dipeptide-repeat pathology in the dentate nucleus of a brain affected by C9 ALS-FTD (bottom right).

in people with sporadic ALS, although they are rare.

Genetic variants that enhance susceptibility to ALS or that modify the clinical phenotype are of immense interest, even if the variants themselves do not cause ALS. For example, large expansions of repeats of the trinucleotide CAG in the coding sequence of the gene *ATXN2* cause spinocerebellar ataxia type 2, in which motor weakness is sometimes an early presentation. It is striking therefore that modest expansions to 27–33 CAG repeats in *ATXN2* increase the risk of developing ALS<sup>5</sup>. By contrast, variants that reduce expression of the axonal guidance gene *EPHA4* improve the overall survival of people with ALS<sup>6</sup>.

### The pathology of ALS

Autopsies of people with ALS reveal the degeneration of motor neurons in the motor cortex of the brain, in the brainstem motor nuclei and in the anterior horns of the spinal cord. As spinal motor neurons degenerate, their target muscles become atrophied. Degeneration of the spinal processes of corticospinal neurons results in scarring in the lateral tracts of the spinal cord. As ALS progresses further, affected

spinal motor neurons shrink and accumulate rounded or thread-like deposits of aggregated proteins that are referred to collectively as inclusions (Fig. 1b). The cytoplasmic inclusions in ALS often become ubiquitinated; an initial target for ubiquitination is TAR DNA-binding protein 43 (TDP-43), encoded by the gene *TARDBP*, which forms the main component of ubiquitinated inclusions in most cases of ALS<sup>7</sup>.

Other pathological features are associated with specific genes. For example, cases of ALS caused by a large expansion of a hexanucleotide repeat in *C9orf72* show intranuclear RNA foci<sup>8</sup>, as well as neuronal cytoplasmic inclusions, predominantly in the cerebellum and hippocampus, that contain the protein sequestosome-1 (also known as ubiquitin-binding protein p62 and encoded by the gene *SQSTM1*) but are distinct from TDP-43 inclusions that are also present in such individuals (ref. 9). Cases of ALS caused by mutations in the genes *SOD1* or *FUS* are pathologically distinct because they exhibit inclusions of abnormal SOD1 or FUS proteins, respectively, rather than those of TDP-43. In addition to these findings in motor neurons, there is also abundant evidence of relevant pathology in non-neural cell types (for example, insidious astrogliosis and microgliosis). It is probable that both of these forms of non-cell-autonomous cellular reactivity influence adversely the progression of ALS.

### Pathogenic mechanisms of ALS

The molecular era of discovery in ALS began with the identification of dominant mutations in the gene *SOD1*, which encodes an abundant, ubiquitously expressed cytoplasmic enzyme called Cu–Zn superoxide dismutase<sup>2</sup>. An important antioxidant, the normal function of SOD1 is to catalyse the conversion of highly reactive superoxide (most frequently produced by errors in mitochondria) to hydrogen peroxide or oxygen.

The expression of mutant *SOD1* in mice demonstrated that the degeneration of motor neurons is driven by one or more acquired toxicities of the mutant protein<sup>10,11</sup> and is independent of dismutase activity<sup>12</sup>. The more than 170 ALS-causing mutations that have now been identified (<http://alsod.iop.kcl.ac.uk/>) lie in almost every region of the 153-amino-acid SOD1 polypeptide. Moreover, although many variants retain partial or full dismutase activity, there is no correlation between a reduction in activity and the age of disease onset or the speed of disease progression<sup>13</sup>. These findings led to the consensus that disease arises from one or more toxic properties of the many SOD1 mutants rather than from reduced dismutase activity.

A sobering reality, however, is that in the 23 years since the discovery of mutations in *SOD1*, no consensus on the main toxicity of mutant SOD1 has emerged. Instead, a plethora of toxic mechanisms that mediate the degeneration and death of motor neurons have been proposed (Fig. 2). A prominent finding is that a proportion of each ALS-causing SOD1 mutant fails to fold properly, which implicates the accumulation of misfolded SOD1 as a possible contributor to toxicity in ALS. Misfolded SOD1 forms ubiquitinated cytoplasmic inclusions that can occur early in ALS and that escalate as the disease progresses<sup>14</sup>.

The accrual of ubiquitinated SOD1 aggregates in people with *SOD1* mutations is paralleled by the accrual of ubiquitinated TDP-43 aggregates in people with *TARDBP* mutations (as well as in most people with sporadic ALS), which highlights a correlation between protein aggregation and ALS. However, as has been demonstrated for other neurodegenerative diseases, large aggregates of disease-causing mutant SOD1 are not sufficient to drive disease because their elimination fails to affect any aspect of the fatal disease that develops in mice expressing ALS-linked mutants of SOD1 (ref. 15).

### Non-cell-autonomous toxicity

Similar to the genes implicated in other main neurodegenerative diseases, all genes in which ALS-causing mutations occur are expressed in many cell types. Indeed, it is now clear that ALS arises, in part, through non-cell-autonomous mechanisms. This means that the disease is the result of a combination of damage from mutant SOD1 in both motor neurons and their glial partners, rather than from damage to neurons alone.

For mutant SOD1, this concept is underscored by studies in mice



revealing that high levels of mutant SOD1 expression in all motor neurons is not sufficient to cause early onset disease<sup>16</sup>. Conversely, a reduction in the synthesis of mutant SOD1 in motor neurons does not slow the rate of progression after the onset of disease, even when applied before symptoms occur<sup>17–19</sup>. Therefore, ALS is a disease not just of the motor neuron but also of the motor system, which is comprised of motor neurons and intimately associated cells of several types.

### The crucial role of glia in ALS

The importance of glial cells in the degeneration and death of motor neurons emerged from studies in which the synthesis of mutant SOD1 was silenced in microglia, astrocytes or oligodendrocyte precursor cells. Microglia, which are the innate immune cells of the nervous system, become activated in all types of ALS (Fig. 2b). The synthesis of mutant SOD1 by microglia is an important determinant of rapid disease progression, as determined by selectively silencing the mutant gene *SOD1* in microglia<sup>19</sup> or by using cell grafts to replace microglia expressing mutant SOD1 with normal microglia<sup>20</sup>. Consistent with these findings, inhibition of the transcription factor NF- $\kappa$ B suppresses this neuroinflammatory component of microglial toxicity in co-cultured motor neurons<sup>21</sup>.

A further mechanism of damage that results from mutant SOD1 produced by microglia is counterintuitive: stimulation of the excessive extracellular production of superoxide<sup>22</sup>. Misfolded mutant SOD1 can associate with the small GTPase RAC1, which controls the activation of NADPH oxidase, a complex that produces superoxide (Fig. 2b). So, instead of its normal function of removing intracellular superoxide, mutant SOD1 could drive microglia to produce high levels of extracellular superoxide.

Disturbances in microglial function have also emerged as a potential contributor to ALS that is associated with mutations in *C9orf72*. Recognition that mutations in *C9orf72* result in the decreased expression of *C9orf72* in people with ALS<sup>8</sup> suggests that the loss of *C9orf72* function might contribute to disease. The protein that *C9orf72* encodes is a potential guanine exchange factor for one or more as-yet-unidentified G proteins. Its inactivation in mice results in abnormal microglia and age-related neuroinflammation, providing evidence that non-cell-autonomous, microglial-mediated inflammation might contribute to ALS<sup>23–25</sup>.

A crucial contribution of mutant SOD1 to pathogenesis is driven by oligodendrocytes, which are cells that myelinate the axons of upper motor neurons and the initial axonal segments of lower motor neurons. A reduction in the synthesis of mutant SOD1 early in oligodendrocyte maturation produces a more striking delay in the onset of disease<sup>26</sup> than does similar suppression of mutant SOD1 synthesis in motor neurons<sup>16,19</sup>. Oligodendrocytes also support motor neuron function by directly supplying the energy metabolite lactate to the axon through the action of monocarboxylate transporter 1 (MCT 1) (Fig. 2c). Mutant SOD1 impairs the expression of MCT 1 by oligodendrocytes in mouse models of ALS<sup>27</sup>. A similar reduction in the accumulation of MCT 1 is found in sporadic ALS<sup>27</sup>, which is consistent with a non-cell-autonomous role for the reduced supply of energy from oligodendrocytes as a general component of ALS pathogenesis.

Another type of glial cell, the astrocyte, provides motor neurons with nutrients, ion buffering and recycling of the neurotransmitter glutamate. The selective reduction of mutant SOD1 synthesis by astrocytes in mice slowed the onset<sup>28</sup> or progression<sup>18</sup> of disease. This delay was accompanied by a delay in the activation of microglia, demonstrating a functional crosstalk between mutant-SOD1-expressing astrocytes and microglia.

One of the earliest proposed mechanisms to underlie ALS was glutamate excitotoxicity, which is the excessive firing of motor neurons that is derived from a failure to rapidly remove synaptic glutamate (Fig. 2d). Astrocytes limit the firing of motor neurons through the swift recovery of glutamate, a function that is mediated by excitatory amino acid transporter 2 (EAAT2), which transports glutamate into the astrocyte (Fig. 2d). The loss of EAAT2 has been observed both in SOD1-mutant rodent models of ALS<sup>14,29</sup> and in samples from people with familial or sporadic ALS<sup>30</sup>. The resulting failure of astrocytes to quickly clear

### BOX 1

## ALS overlap syndromes

Careful observation of people with ALS over the past 30 years has revealed clinical, pathological and genetic overlap with several other neurodegenerative disorders. In particular, the loss of motor neurons may be accompanied by the loss of cortical neurons in the frontal and temporal cortices of the brain. This correlates clinically with FTD, a condition of impaired judgment and executive skills, which often leads to behavioural disturbances. About 20% of people with ALS meet the clinical criteria for a concomitant diagnosis of FTD, although as many as 50% of people with ALS experience cognitive impairment. Less frequently, ALS occurs together with Paget's disease of bone (PDB) or inclusion body myopathy (IBM). Similar to ALS, both FTD and IBM are characterized by inclusions of TDP-43 and related RNA-binding proteins. The relationship between ALS, FTD, PDB and IBM has been extended though genetic evidence, which potentially places them in the continuum of a broader degenerative disorder.

synaptic glutamate triggers the repetitive firing of action potentials and a corresponding increase in calcium influx, as well as endoplasmic reticulum (ER) and mitochondrial stress as the result of overwhelming the calcium storage capacities of these organelles.

Astrocytes also protect motor neurons from excitotoxic damage through the release of an unidentified soluble factor or factors that induce motor neurons to upregulate the glutamate receptor subunit GluR-2 (ref. 31). The incorporation of GluR-2 into glutamate receptors in neurons reduces the permeability of these receptors to calcium, which provides protection from excitotoxicity by decreasing the influx of calcium. Astrocytes that express mutant SOD1 fail to regulate GluR-2 expression in co-cultured neurons, thereby increasing their vulnerability to excitotoxic damage<sup>31</sup>.

Several teams of researchers have used *in vitro* co-cultures of motor neurons and astrocytes (or astrocyte-conditioned medium) to show that astrocytes expressing ALS-linked mutations produce a toxicity that diffuses to motor neurons<sup>32–36</sup> (Fig. 2d). However, there is no consensus on the identity of the toxic species. Notably, astrocytes from people with familial ALS or sporadic ALS (obtained directly from autopsy samples<sup>36</sup> or by isolating neuronal precursor cells that can be converted into astrocytic precursor cells and then astrocytes<sup>35</sup>) are toxic to co-cultured normal motor neurons. This finding<sup>35</sup> is especially provocative because it indicates that neuronal precursor cells in portions of tissue from people with sporadic ALS have already acquired damage and that this damage is retained following several divisions of these cells in culture and their subsequent differentiation. Whether toxicity from sporadic ALS-derived astrocytes is mediated by changes in SOD1 (ref. 35) or not<sup>36</sup> remains unsettled.

Most importantly, a non-cell-autonomous contribution of astrocytes to ALS-like disease has been demonstrated in rodents expressing mutant SOD1 in which transplantation to the spinal cord of lineage-restricted astrocyte precursors without SOD1 mutations delayed progression of the disease<sup>37</sup>.

### ALS genes induce ER stress or impair protein degradation

ER stress has been implicated broadly in ALS (Fig. 2a). Initial evidence arose from studies of mutant SOD1 in which misfolded SOD1 binds to the cytoplasmic surface of the ER integral membrane protein derlin-1 (ref. 38). This binding leads to the inhibition of ER-associated degradation (ERAD), the pathway for extraction and degradation of misfolded proteins from the ER. Moreover, relieving ER stress delays the progression of disease in an animal model of ALS<sup>39</sup>.

There is now overwhelming evidence to show that disruption of the

two main protein clearance pathways, the ubiquitin–proteasome system and autophagy, can be central components of the disease mechanism in ALS (Fig. 2a). Several ALS-causing mutations occur in genes with products that are involved directly in protein degradation, including ubiquilin-2 (ref. 40) and sequestosome-1 (ref. 41), both of which function as adapters that bring polyubiquitinated proteins to the proteasome or the autophagosome for degradation. Mutations have also been reported in optineurin<sup>42</sup>, a proposed receptor for autophagy<sup>43</sup>, and valosin-containing protein (VCP)<sup>44</sup>, which has a role in ERAD and sorting endosomal proteins. Other studies have reported FTD-linked mutations in *CHMP2B*<sup>45</sup>, which encodes a protein that has been implicated in maturation of the autophagosome and endosomal cargo sorting and degradation. ALS-linked mutations are also found in *VAPB*<sup>46</sup>, the product of which functions in the unfolded protein response in the delivery of ER-ejected substrates to the proteasome. A preponderance of biochemical evidence has demonstrated a decrease in the activities of the proteasome in lumbar spinal cords before symptoms occur in mice that express mutant SOD1 (ref. 47) or following the sustained expression of mutant SOD1 in a cultured line of neurons<sup>48</sup>.

### Axonal disorganization and disrupted transport in ALS

Disorganization of the axonal cytoskeleton, and especially of the neurofilaments, is a conspicuous feature both of familial ALS and sporadic ALS (Fig. 2e). As the most asymmetric cells in nature, and with axons that can reach more than 1 metre in length, motor neurons must rely on axonal transport to deliver components that are synthesized in the cell bodies to axons and synapses. ALS-linked mutant SOD1 has been demonstrated to slow both anterograde<sup>49</sup> and retrograde<sup>50,51</sup> transport routes months before neurodegeneration. Indeed, reduction in retrograde transport through mutations in dynactin<sup>52</sup>, which is an activator of the retrograde motor cytoplasmic dynein<sup>53</sup>, provokes motor neuron disease in humans.

Owing to the peculiar architecture of neurons, it is a challenge for these cells to alter local gene expression at the synapse in response to neuronal input or changes in the synaptic environment. To achieve this, neurons must transport all necessary components for translation (for example, messenger RNA, ribosomes and translation factors) to distal sites for local protein synthesis<sup>54</sup>. The spatial distribution of mRNAs depends on the

proper microtubule-dependent transport of neuronal RNA transport granules and other factors, and it is regulated by several RNA-binding proteins that are associated with ALS, including TDP-43, FUS and heterogeneous nuclear ribonucleoprotein (hnRNP) A1. ALS-causing mutations in TDP-43 impair the axonal transport of RNA granules in *Drosophila* and in cultured neurons, including motor neurons derived from people with ALS<sup>55</sup>.

### A prion-like spread in inherited ALS

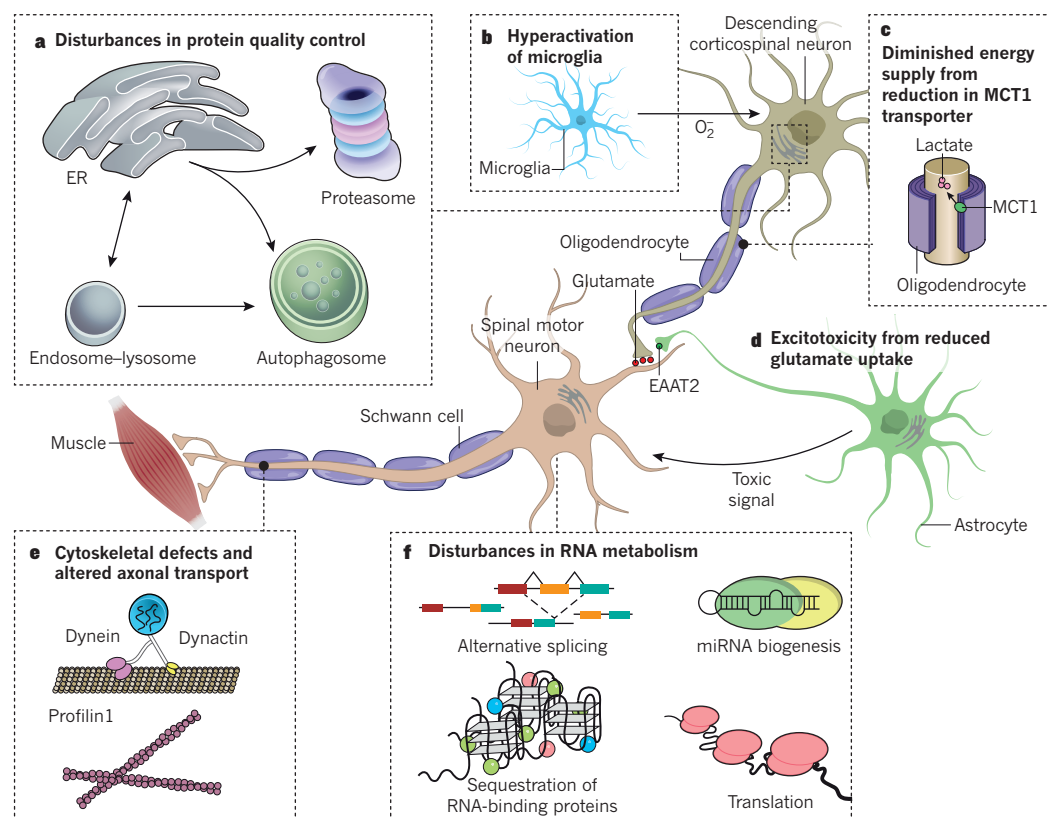
The prion-like, templated conversion of a natively folded protein into a misfolded version of itself is now recognized as a prominent feature of the cell-to-cell spread of protein aggregates in neurodegenerative diseases. Examples include  $\alpha$ -synuclein templating in Parkinson's disease, amyloid- $\beta$  aggregation in Alzheimer's disease and tau misfolding in chronic brain injury (reviewed in further detail in ref. 56). Evidence for similar templated toxicity has emerged for misfolded SOD1 (refs 57 and 58), with wild-type SOD1 exacerbating the toxicity of mutant SOD1 in mice<sup>59</sup>. Prion-like propagation and development of disease that is initiated focally has been shown to occur after the injection of lysates containing mutant SOD1 into mice that express mutant SOD1 (ref. 60). This finding replicates the correlation between focal initiation and spreading in people with familial ALS or sporadic ALS<sup>61</sup>.

That said, prion-like propagation of SOD1 (or other ALS-linked proteins) has not been achieved in rodents without the coexistence of a pre-existing, weakly active mutant ALS gene. It is unresolved whether this evidence challenges the prion-like spread model of sporadic ALS or, alternatively, whether it raises the possibility that there must be a pre-existing sensitivity in individuals who develop sporadic ALS that facilitates such spreading. Coupled with the recognition that misfolded mutant SOD1 can be secreted by motor neurons or astrocytes<sup>62</sup>, potentially through the newly discovered pathway in which misfolded proteins are secreted unconventionally as an adaptation to proteasome dysfunction<sup>63</sup>, stochastic focal initiation provides a plausible mechanism for the age-dependent onset of disease and its subsequent spread. As most cases of ALS are marked by aggregated TDP-43 rather than SOD1, an unresolved question is whether TDP-43 also exhibits templated misfolding that can spread from cell to cell.

**Table 1 | The genetics of ALS**

Locus	Gene	Protein	Protein function	Mutations	Proportion of ALS		Date of discovery
					Familial	Sporadic	
21q22.1	<i>SOD1</i>	Cu–Zn superoxide dismutase	Superoxide dismutase	>150	20%	2%	1993 (ref. 2)
2p13	<i>DCTN1</i>	Dynactin subunit 1	Component of dynein motor complex	10	1%	<1%	2003 (ref. 52)
14q11	<i>ANG</i>	Angiogenin	Ribonuclease	>10	<1%	<1%	2006 (ref. 141)
q36	<i>TARDBP</i>	TDP-43	RNA-binding protein	>40	5%	<1%	2008 (refs 67 and 142)
16p11.2	<i>FUS</i>	FUS	RNA-binding protein	>40	5%	<1%	2009 (refs 68 and 69)
9p13.3	<i>VCP</i>	Transitional endoplasmic reticulum ATPase	Ubiquitin segregase	5	1–2%	<1%	2010 (ref. 44)
10p15-p14	<i>OPTN</i>	Optineurin	Autophagy adaptor	1	4%	<1%	2010 (ref. 42)
9p21-22	<i>C9orf72</i>	C9orf72	Possible guanine nucleotide exchange factor	Intronic GGGGCC repeat	25%	10%	2011 (refs 8 and 77)
Xp11.23-Xp13.1	<i>UBQLN2</i>	Ubiquilin 2	Autophagy adaptor	5	<1%	<1%	2011 (ref. 40)
5q35	<i>SQSTM1</i>	Sequestosome 1	Autophagy adaptor	10	<1%	?	2011 (refs 41 and 143)
17p13.2	<i>PFN1</i>	Profilin-1	Actin-binding protein	5	<1%	<1%	2012 (ref. 144)
12q13.1	<i>HNRNPA1</i>	hnRNP A1	RNA-binding protein	3	<1%	<1%	2013 (refs 70 and 71)
5q31.2	<i>MATR3</i>	Matrin 3	RNA-binding protein	4	<1%	<1%	2014 (ref. 76)
2q36.1	<i>TUBA4A</i>	Tubulin $\alpha$ -4A chain	Microtubule subunit	7	<1%	<1%	2014 (ref. 145)
22q11.23	<i>CHCHD10</i>	Coiled-coil-helix-coiled-coil-helix domain-containing protein 10	Mitochondrial protein of unknown function	2	<1%	<1%	2014 (ref. 146)
12q14.1	<i>TBK1</i>	Serine/threonine-protein kinase TBK1	Regulates autophagy and inflammation	10	?	?	2015 (ref. 147)





**Figure 2 | Mechanisms of disease implicated in ALS.** **a**, Familial ALS-associated mutations frequently affect genes that are components of the cellular protein quality control system. Other mutations, such as those in *SOD1*, affect protein folding. **b**, Hyperactivation of microglia produces extracellular superoxide, which triggers inflammation and degeneration in motor neurons. **c**, A reduction in the levels of the lactate transporter MCT 1 diminishes energy supplied by oligodendrocytes to motor neurons. **d**, A failure of astrocytes to clear synaptic glutamate via the transporter EAAT2 triggers repetitive firing of motor neurons and excitotoxicity. **e**, Disruption of the cytoskeleton and impaired axonal transport limits the exchange of essential macromolecules and organelles between the neuronal cell body and distal compartments. **f**, Disturbances in aspects of RNA metabolism, including RNA processing, transport and utilization, are largely the result of impaired hnRNP function.

## The intersection of RNA biology and ALS pathogenesis

In 2006, Virginia Lee and colleagues reported the mislocalization of RNA-binding protein TDP-43 from its predominantly nuclear location to ubiquitin-containing cytoplasmic inclusions in affected areas of the brain and the spinal cord of people with ALS<sup>7</sup>. TDP-43 mislocalization is now recognized widely as the hallmark of both sporadic ALS and most familial forms of ALS. This seminal discovery has implications beyond ALS because TDP-43 mislocalization to cytoplasmic inclusions is also the hallmark of FTD that lacks tau-containing inclusions (about half of all cases of FTD) and inclusion body myopathy (IBM)<sup>7,64</sup>, which are diseases that show genetic overlap with ALS. Moreover, TDP-43 pathology is also found as a secondary pathological feature in a subset of people with Alzheimer's disease or Parkinson's disease<sup>65,66</sup>. The importance of TDP-43 in pathogenesis was cemented by the identification of ALS-causing mutations in this protein<sup>67</sup>. The subsequent identification of ALS-causing mutations in related proteins that bind RNA, including FUS and hnRNP A1, focused substantial attention on the role of RNA biology in ALS pathogenesis<sup>68–71</sup> (Fig. 2f).

TDP-43, FUS and hnRNP A1 are members of the hnRNP family of proteins that regulates RNA metabolism at every stage of the RNA life cycle. They bind to thousands of RNA targets<sup>72–75</sup>, which means that a disturbance in the function of one or more of these proteins has the potential to affect RNA metabolism on a broad scale. Further links between the pathogenesis of ALS and RNA biology have emerged from the identification of ALS-causing mutations in the RNA-binding protein matrin-3 (ref. 76), an appreciation of the increased risk of developing ALS in association with certain alleles of the RNA-binding protein ataxin-2 (ref. 5) and the recognition of RNA-related mechanisms of disease that are associated with mutations in *C9orf72* (refs 8 and 77).

## Phase separation gives rise to membraneless organelles

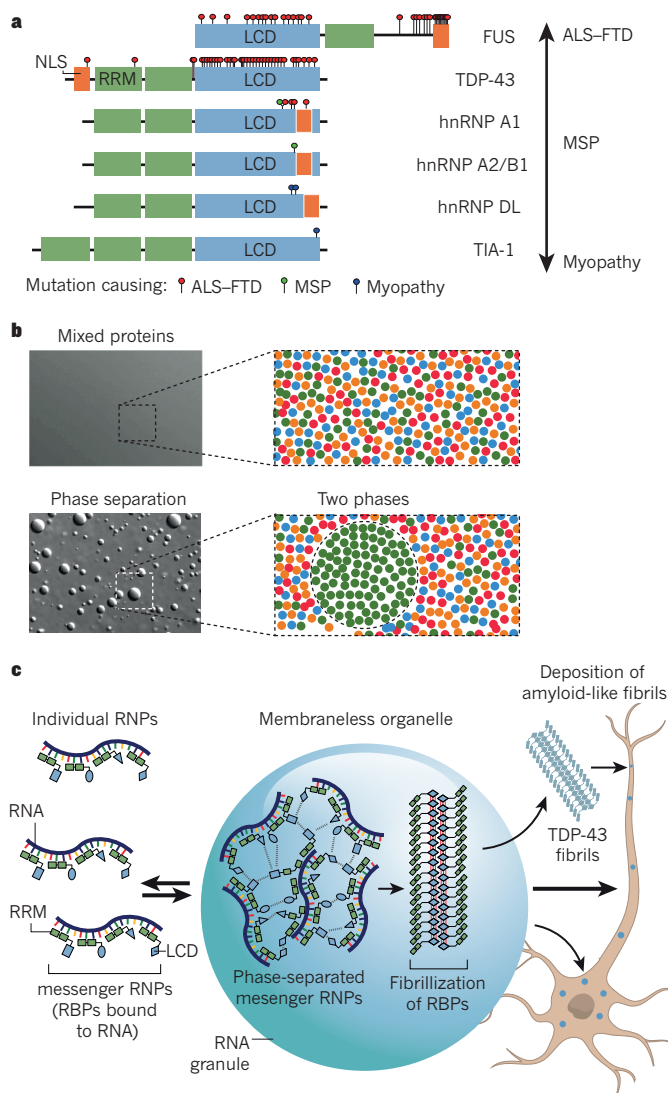
RNA metabolism occurs in complex RNA–protein assemblies that can coalesce into a variety of membraneless organelles such as nucleoli and stress granules. Interestingly, these organelles behave as complex liquids that arise through phase separation, a process in which protein-laden RNAs separate from the surrounding aqueous nucleoplasm or cytoplasm

in a manner that is akin to the separation of oil from vinegar<sup>78</sup>. Phase separation is mediated by low-complexity domains that are present in RNA-binding proteins such as TDP-43, FUS and hnRNP A1 (refs 79–81) (Fig. 3a, b). The assembly of membraneless organelles is a strategy of cellular compartmentalization that governs many biological processes. However, the contribution of phase transition to this process presents a risk because RNA-binding proteins with low-complexity domains, which are prone to fibrillization, are placed in close proximity. Indeed, mutations that cause ALS are found frequently in the low-complexity domains of TDP-43 (ref. 82), FUS<sup>83</sup> and hnRNP A1 (ref. 70). As a consequence, these mutations alter the dynamics of membraneless organelles and also accelerate fibrillization, which results in the formation of amyloid-like fibrils that are deposited in the cell bodies and the neuropil<sup>79–81,84</sup> (Fig. 3c).

Mutations in the low-complexity domains of at least six different hnRNPs result in a clinico-pathological spectrum that ranges from ALS and FTD to IBM (Fig. 3a and Box 1). Notably, some disease-causing mutations in RNA-binding proteins do not affect low-complexity domains. For example, several ALS-causing mutations in FUS and hnRNP A1 disturb the nuclear localization sequence of these proteins and result in their accumulation in the cytoplasm<sup>68,69,71</sup>. Phase transition by RNA-binding proteins that contain low-complexity domains is exquisitely dependent on concentration<sup>79</sup>, and it is probable that the increased accumulation of FUS and hnRNP A1 in the cytoplasm as a consequence of mutations that affect the nuclear localization signals of these proteins is sufficient to drive excess phase separation, as shown by the hyperassembly of stress granules in cells derived from people with relevant mutations<sup>70,85</sup>.

## RNA metabolism defects in ALS

Disturbance of the normal phase transitions carried out by RNA-binding proteins can have deleterious consequences, including altering the material properties of RNA granules and impairing their function<sup>55</sup>. Moreover, persistent assembly of RNA-binding proteins in the highly concentrated liquid state may promote the formation of amyloid-like fibrils that have toxic properties<sup>79,80</sup> and may also result in a partial or complete loss of the normal function of important RNA-binding proteins<sup>86</sup>. A well known feature of ALS histopathology is the redistribution of TDP-43 from the



**Figure 3 | ALS mutations impair the assembly, dynamics and function of membraneless organelles.** **a**, A schematic representation of six hnRNPs (FUS, TDP-43, hnRNP A1, hnRNP A2/B1, hnRNP DL and TIA-1) harbouring mutations that produce a spectrum of disease that ranges from ALS or FTD to myopathy. LCD, low-complexity domain; MSP, multisystem proteinopathy; NLS, nuclear localization signal; RRM, RNA recognition motif. **b**, RNA-binding proteins that contain a low-complexity domain can undergo phase separation, which is the transition from a single, mixed phase (top) to two distinct phases (bottom), one of which is a concentrated liquid droplet (green). **c**, Phase separation contributes to the assembly, dynamics and liquid properties of membraneless organelles; however, the high concentration and close positioning of the low-complexity domains risks the transitioning of such proteins (for example, TDP-43) to pathological, amyloid-like fibrils. RBP, RNA-binding protein; RNP, ribonucleoprotein.

nucleus to the cytoplasm<sup>7</sup>. A similar redistribution is observed for FUS and hnRNP A1 when disease-causing mutations occur in the genes that encode these proteins<sup>68,69,71</sup>. This redistribution of proteins might reflect a cytoplasmic sink that is produced by the hyperassembly of cytoplasmic granules or by poorly dynamic RNA granules that fail to disassemble appropriately, the deposition of amyloid-like fibrils and defects in nucleocytoplasmic trafficking, as well as other potential mechanisms.

The culmination of this redistribution is the depletion of RNA-binding proteins in the nucleus that has the potential to cause a considerable loss of nuclear function. A well known function of TDP-43 in the nucleus is the regulation of alternative splicing<sup>87</sup>. Experimental depletion of TDP-43 in rodents was found to alter hundreds of splicing events in the brain, resulting in the depletion of several RNAs that encode

synaptic proteins<sup>73</sup>. The loss of nuclear TDP-43 also facilitates the use of cryptic splice sites<sup>88</sup> that, in general, might lower the levels of correctly spliced protein-encoding mRNAs. Furthermore, TDP-43 autoregulates its synthesis<sup>73</sup>, which establishes the possibility of a feed-forward mechanism that amplifies the impact of the partial loss of TDP-43 function.

The loss of FUS or hnRNP A1 from the adult nervous system produces defects analogous to those associated with the loss of TDP-43, although different subsets of mRNAs are linked to the depletion of each of these RNA-binding proteins<sup>74,75</sup>. An important, unanswered question concerns the extent to which ALS caused by other genetic perturbations, especially *C9orf72*-related ALS, also involves disturbances in RNA biology that intersect mechanistically with mutations in TDP-43, FUS and hnRNP A1.

### The biogenesis of regulatory RNA and its function in ALS

Both TDP-43 and FUS are components of macromolecular complexes that generate small non-coding RNAs known as microRNAs (miRNAs) with functions in RNA silencing. The loss of TDP-43 or FUS results in a reduction in the expression of miRNAs in model systems, including *Drosophila* models and induced pluripotent stem (iPS)-derived motor neurons from people with TDP-43 mutations, which suggests a possible role for altered RNA silencing in ALS<sup>89</sup>. Various miRNAs contribute to the maintenance of neuromuscular junctions, implying that motor neurons might be particularly sensitive to disturbances in the biogenesis of miRNA<sup>90,91</sup>. Indeed, global downregulation of miRNAs has been reported in motor neurons from people with sporadic ALS<sup>92</sup>, although the role of reduced levels of miRNAs in the pathogenesis of ALS remains to be established<sup>92</sup>. Nonetheless, the expression of regulatory RNAs seems to be altered robustly and consistently in the serum of people with ALS, and this could present an opportunity for the development of biomarkers<sup>93</sup>.

### The curious case of *C9orf72*-related ALS and FTD

Although the identification of ALS-causing mutations that affect SOD1 and RNA-binding proteins highlighted pathophysiological pathways through which disease might arise, most of the genetic burden of ALS remained unaccounted for until 2011. Genetic linkage studies<sup>94,95</sup> followed by several large genome-wide association studies<sup>96–99</sup> identified the location of a gene in the chromosome 9p21 locus in which mutations cause both ALS and FTD. During sequencing of the non-coding regions of candidate genes in chromosome 9p21, a pathogenic expansion of a hexanucleotide repeat in *C9orf72* was identified as the basis for *C9orf72*-related ALS and FTD (C9 ALS-FTD)<sup>8</sup>. In healthy individuals, the sequence GGGGCC was present as 2–23 repeats but in affected individuals it was expanded to hundreds or thousands of repeats. In parallel, an independent study also discovered a pathogenic expansion of GGGGCC repeats in *C9orf72* (ref. 77).

### The consequences of repeat expansion in *C9orf72*

Three non-exclusive mechanisms have been proposed through which expanded GGGGCC repeats might cause C9 ALS-FTD (Fig. 4). First, a reduction in the expression levels of *C9orf72* in people with C9 ALS-FTD<sup>100</sup> has led to speculation that the loss of *C9orf72* protein may contribute to disease. The function of *C9orf72* is poorly understood; however, the protein is known to contain a conserved DENN domain and can function as a guanine-nucleotide exchange factor for several Rab proteins in experimental systems<sup>101</sup>. In cultured cells and in zebrafish, the depletion of endogenous *C9orf72* can exacerbate the toxicity of aggregation-prone proteins such as polyglutamine-expanded ataxin-2 (ref. 102). However, the reduction of endogenous *C9orf72* mRNA with antisense oligonucleotides was well tolerated in mice and did not result in impairments to behaviour or motor functions<sup>103</sup>. Furthermore, the conditional knockout of the gene *C9orf72* in the brains of mice did not cause noticeable motor neuron or other neurodegenerative phenotypes, and there was no evidence of the hallmark pathological features of ALS or FTD<sup>104</sup>. Mice in which there was a complete ablation of *C9orf72* in all tissues developed and aged normally without the occurrence of motor neuron disease; however, they also developed progressive splenomegaly



and lymphadenopathy<sup>23–25,105</sup>. These mice were found to have abnormal macrophages and microglia, as well as age-related neuroinflammation<sup>23,105</sup> and signs of autoimmunity<sup>25</sup>, which raises the possibility of a non-cell-autonomous inflammatory contribution to C9 ALS–FTD.

Nevertheless, the dominant inheritance pattern of C9 ALS–FTD, the absence of people with ALS or FTD with null alleles or missense mutations in *C9orf72* and the absence of neurodegeneration in *C9orf72*-knock-out mice provide arguments against the loss of *C9orf72* function as the sole driver of disease. Indeed, most empirical evidence points to the gain of toxic functions as the main mechanisms that drive neurodegeneration in C9 ALS–FTD. For example, the adeno-associated virus-mediated delivery to the brain of a construct that expresses expanded GGGGCC repeats elicits neurodegeneration<sup>106</sup>, although the nature of the toxic species in C9 ALS–FTD remains unclear.

### Gain of toxic function from repeat-containing RNA

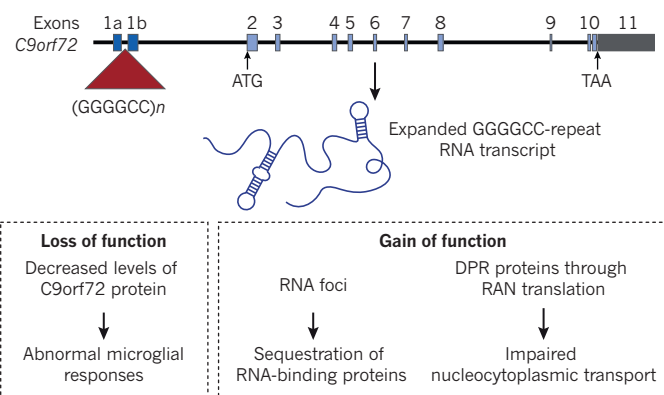
The initial description of the mutation in *C9orf72* was accompanied by evidence to show that RNA foci containing the GGGGCC repeat accumulate in the brains and spinal cords of people with C9 ALS–FTD<sup>8</sup> (Fig. 1b), and this suggested a second possible disease mechanism, involving toxic gain of function by repeat-containing RNA (Fig. 4). It was then noted that the gene *C9orf72* can be transcribed bidirectionally and that foci containing sense (GGGGCC) or antisense (CCCCGG) RNA transcripts accumulate in affected cells<sup>107–109</sup>. The accrual of such foci in C9 ALS–FTD is reminiscent of the pathological RNA foci that are observed in myotonic dystrophy type 1, myotonic dystrophy type 2 and fragile X-associated tremor and ataxia syndrome, which are also caused by the expansion of nucleotide repeats in non-coding regions<sup>110</sup>. In these diseases, the accumulated repeat-containing RNA sequesters RNA-binding proteins that are involved in splicing, which leads to defects in splicing that underlie some aspects of pathogenesis<sup>110</sup>. Similarly, a number of RNA-binding proteins bind to expanded GGGGCC or GGCCCC repeats *in vitro*, and a rare co-localization with RNA foci has been observed for several of these proteins in tissue from affected individuals<sup>111–115</sup>.

Simple model systems have illustrated the functional consequences of the sequestration of some hexanucleotide repeat-binding proteins, including transcriptional activator protein Pur-α and Ran GTPase-activating protein 1 (RanGAP1), but the contribution of these interactions to the development of disease is not yet established<sup>114,116</sup>. Notably, repeats of GGGGCC (but not of CCCCCG) can adopt a stable secondary structure known as a G-quadruplex, which might contribute to the persistence of this species of RNA as well as to its ability to reach distal neurites and associate with RNA-binding proteins in transport granules and potentially interfere with local translation<sup>117–119</sup>.

### Gain of toxic function from dipeptide repeats

Substantial evidence has also accrued to implicate a third disease mechanism in C9 ALS–FTD; specifically, toxicity from DPR proteins that are produced by repeat-associated non-AUG (RAN) translation (Fig. 4). This unconventional type of translation occurs in the absence of an initiating AUG codon and might rely on secondary structures formed by repeat-expanded RNA<sup>120</sup>. In C9 ALS–FTD, RAN translation occurs in all reading frames and from both sense and antisense transcripts, and it results in the production of five DPR proteins: glycine-alanine (GA) and glycine-arginine (GR) from sense GGGGCC transcripts; proline-arginine (PR) and proline-alanine (PA) from antisense GGCCCC transcripts; and glycine-proline (GP) from both sense and antisense transcripts<sup>107–109</sup>. All of these DPR proteins are produced in people with C9 ALS–FTD and they account for the neuronal cytoplasmic and intranuclear inclusions that contain ubiquitin and sequestosome-1 but lack TDP-43 that are found widely in the brain and spinal cord<sup>107,109,113,121,122</sup> (Fig. 1b).

The timing, location and level of expression of each species of DPR protein in the brains of affected people are yet to be clarified. Several reports have described the deposition of DPR proteins in the brains of people with C9 ALS–FTD, and in some instances an inverse relationship has been described between the regional burden of DPR proteins and the



**Figure 4 | Proposed mechanisms for the development of C9 ALS–FTD.** Expansion of an intronic hexanucleotide repeat (GGGGCC) in *C9orf72* from fewer than 23 copies to hundreds or thousands of copies causes C9 ALS–FTD. This mutation results in a modest reduction in the levels of *C9orf72* protein (left) that seems insufficient to cause disease but might contribute to its progression through abnormal microglial responses. Meanwhile, the expression of sense and antisense RNA transcripts that contain the expanded repeat probably drive a toxic gain of function (right). The two main gain-of-function modes that are implicated are: toxicity through the sequestration of RNA-binding proteins in RNA foci by the expanded GGGGCC repeat RNA transcript; and the production of DPR proteins through RAN translation, leading to toxicity through several cellular targets such as membraneless organelles and nuclear pores.

corresponding severity of neurodegeneration<sup>123,124</sup>. These studies were based on the post-mortem examination of brains with end-stage disease and relied on the detection of large inclusions using immunohistochemistry, an approach that probably under-represents the pathological burden of soluble DPR proteins. However, the apparent discrepancy between the burden of DPR-protein deposition, the levels of which are greatest in the cerebellum, and the severity of neurodegeneration, which is greatest in the motor cortex and spinal cord, needs to be resolved to understand the role of DPR proteins in the development of disease.

Some species of DPR proteins have been shown to be toxic in cultured cells and animal models of disease, although high levels of expression were sometimes used to produce short-term toxicity. The arginine-containing DPR proteins GR and PR seem to be most toxic. For example, when GR or PR is added to cells in culture, it enters and accumulates in nucleoli, which leads to defects in RNA processing and subsequent cell death<sup>125</sup>. Similarly, the independent expression of each of the five species of DPR proteins in cultured neurons revealed that GR and PR are very toxic, whereas PA, GA and GP are well tolerated. Observations in *Drosophila* engineered to express each of the five DPR proteins have also shown that GR and PR are extremely toxic to neuronal tissue, whereas GA is modestly toxic and GP and PA seems to be non-toxic<sup>126–128</sup>. A recent discovery is that the arginine-containing DPR proteins GR and PR bind to proteins that contain low-complexity domains<sup>129,130</sup>. Furthermore, GR and PR alter the phase separation of such proteins, resulting in the perturbed assembly, dynamics and function of membraneless organelles such as stress granules and nucleoli<sup>129</sup>. This finding mirrors the defects in phase transitions that are observed with disease-causing mutations in the low-complexity domains of TDP-43, FUS and hnRNP1, suggesting a common pathological mechanism.

However, other investigations have reported that toxicity is associated with the expression of GA in cell culture<sup>131–133</sup> and its adeno-associated virus-mediated delivery to the mouse brain<sup>133</sup>. It should be noted that these efforts to model the toxicity of DPR proteins have used short (fewer than 100) repeats. How the properties of those short DPR proteins compare with the possibly larger products of RAN translation in affected individuals is also unknown.

### A defect in nucleocytoplasmic trafficking

Whereas the nature of the gain of toxic function is still an open question,

converging evidence suggests that impaired nucleocytoplasmic trafficking is one of the downstream consequences of mutations in *C9orf72*. A comprehensive, unbiased screen in *Drosophila* for genetic modifiers of the toxicity that is mediated by expanded GGGGCC repeats identified 18 genes that are connected to the nuclear pore complex and nucleocytoplasmic trafficking<sup>128</sup>. A separate unbiased screen for genetic modifiers of PR toxicity in yeast also identified numerous genes that encode components of the nuclear pore complex and effectors of nucleocytoplasmic trafficking<sup>134</sup>. A third study focused on the nucleocytoplasmic transport factor RanGAP1, which binds to the RNA sequence GGGGCC. Genes encoding RanGAP1 and other nucleocytoplasmic transport factors were identified as modifiers of toxicity mediated by expanded GGGGCC repeats in *Drosophila*<sup>116</sup>. Consistent with these results, morphological abnormalities were found in the nuclear envelope architecture in both cell-based and animal models of disease, as well as in the brains of people with C9 ALS–FTD. Moreover, defects in the nucleocytoplasmic transport of RNA and proteins were found in neurons derived from the iPSCs of people with C9 ALS–FTD<sup>116,128,134</sup>.

### Approaches to therapy for C9 ALS–FTD

The relative contributions of the various proposed modes of toxicity to the development of C9 ALS–FTD is an important consideration that will influence strategies for therapeutic intervention. Efforts are underway to impede RAN translation with small molecules but the success of such an approach will depend on the role of DPR proteins in disease. Irrespective of the main basis for the toxic gain of function, the mutant gene *C9orf72* presents an attractive target for therapeutic intervention. For example, antisense oligonucleotides are able to reverse pathological features in neurons derived from iPSCs<sup>103,135,136</sup> or in fibroblasts<sup>103</sup> from people with C9 ALS–FTD. Indeed, neurons and glial cells derived from iPSCs might prove to be a useful model system in which to develop approaches for mitigating toxicity related to mutant *C9orf72* even before the basis of toxicity has been elucidated fully.

Therapeutic efforts will be aided further by the development of transgenic mouse models that express human *C9orf72* that contain about 450 hexanucleotide repeats, which recapitulate aspects of the molecular pathology<sup>24,137–139</sup>, neuropsychological deficits<sup>24,139</sup> and the motor phenotype<sup>139</sup> of C9 ALS–FTD. It is also particularly promising that pathological abnormalities can be reversed, and that the development of neuropsychological deficits can be delayed, by a single-dose infusion of an antisense oligonucleotide that induces the catalytic degradation of hexanucleotide-containing RNAs without exacerbating a reduction in RNAs encoding the *C9orf72* protein<sup>24</sup>.

### Looking forward

Clearly, there has been dramatic progress towards defining the genetic topography and molecular biology of ALS. There is also little doubt that the pace of discovery will continue or even accelerate in several areas of research.

First, it is certain that our understanding of the genetic basis of ALS will continue to evolve. Research programmes are already in place to collect and sequence thousands of whole genomes from people with ALS. More genes that are implicated in ALS are likely to be defined, both through conventional Mendelian genetics and through enhanced association studies that identify increased burdens of rare genetic variants, including those found in non-coding DNA. In parallel, enhanced scoring and recording of quantifiable clinical parameters will permit the definition of variants that modify the phenotype of ALS. The existence of extensive ALS genome databases will enable the first comprehensive studies of epistasis, characterizing the interactions of numerous genes to perturb the viability of motor neurons.

Second, although the past two decades have witnessed extraordinary progress in understanding familial ALS, it is probable that insights that help to elucidate sporadic ALS will be acquired. One view is that all cases of sporadic ALS will ultimately be shown to reflect several genetic determinants. Alternatively, there is increasing interest in exogenous factors

that might trigger sporadic neurodegeneration, and atypical infections or the activation of endogenous retroviruses<sup>140</sup> are proposed to have such a role. Although the role of external environmental factors in ALS has been elusive, there is fresh interest in the influence of the intrinsic environment, represented by the microbiome, on development of the disease.

Last, and perhaps most importantly, there will be considerable achievements in the development of therapies for ALS. Although daunting, the complexity of the molecular pathology of ALS is promising as a roadmap for defining therapeutic targets. Moreover, for types of ALS that arise from well-defined genetic defects, advances in gene silencing and gene editing technologies will permit personalized therapeutic programmes. When combined with improved methods for the delivery of therapies to the central nervous system, these approaches will lead to strategies for attenuating the lethal course of ALS. ■

Received 7 July; accepted 13 September 2016.

- Al-Chalabi, A. & Hardiman, O. The epidemiology of ALS: a conspiracy of genes, environment and time. *Nature Rev. Neurol.* **9**, 617–628 (2013).
- Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59–62 (1993).  
**The first report of a gene defect that causes ALS.**
- Gamez, J. *et al.* Mutational analysis of the Cu/Zn superoxide dismutase gene in a Catalan ALS population: should all sporadic ALS cases also be screened for SOD1? *J. Neurol. Sci.* **247**, 21–28 (2006).
- Cooper-Knock, J. *et al.* Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. *Brain* **135**, 751–764 (2012).
- Elden, A. C. *et al.* Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* **466**, 1069–1075 (2010).  
**The first report of a relationship between microsatellite expansion in ATXN2 and susceptibility to ALS.**
- Van Hoecke, A. *et al.* EPHA4 is a disease modifier of amyotrophic lateral sclerosis in animal models and in humans. *Nature Med.* **18**, 1418–1422 (2012).
- Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
- DeJesus-Hernandez, M. *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245–256 (2011).  
**This paper and ref. 79 were the first reports to describe pathogenic microsatellite expansion in C9orf72.**
- Al-Sarraj, S. *et al.* p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTL and MND/ALS. *Acta Neuropathol.* **122**, 691–702 (2011).
- Gurney, M. E. *et al.* Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation. *Science* **264**, 1772–1775 (1994).
- Wong, P. C. *et al.* An adverse property of a familial ALS-linked SOD1 mutation causes motor neuron disease characterized by vacuolar degeneration of mitochondria. *Neuron* **14**, 1105–1116 (1995).
- Brujin, L. I. *et al.* Aggregation and motor neuron toxicity of an ALS-linked SOD1 mutant independent from wild-type SOD1. *Science* **281**, 1851–1854 (1998).  
**This study demonstrated that ALS-causing mutant SOD1 generates a toxicity that is independent of its dismutase activity.**
- Cleveland, D. W., Laing, N., Hulse, P. V. & Brown, R. H. Jr. Toxic mutants in Charcot's sclerosis. *Nature* **378**, 342–343 (1995).
- Brujin, L. I. *et al.* ALS-linked SOD1 mutant G85R mediates damage to astrocytes and promotes rapidly progressive disease with SOD1-containing inclusions. *Neuron* **18**, 327–338 (1997).
- Parone, P. A. *et al.* Enhancing mitochondrial calcium buffering capacity reduces aggregation of misfolded SOD1 and motor neuron cell death without extending survival in mouse models of inherited amyotrophic lateral sclerosis. *J. Neurosci.* **33**, 4657–4671 (2013).
- Yamanaka, K. *et al.* Mutant SOD1 in cell types other than motor neurons and oligodendrocytes accelerates onset of disease in ALS mice. *Proc. Natl Acad. Sci. USA* **105**, 7594–7599 (2008).
- Ralph, G. S. *et al.* Silencing mutant SOD1 using RNAi protects against neurodegeneration and extends survival in an ALS model. *Nature Med.* **11**, 429–433 (2005).
- Yamanaka, K. *et al.* Astrocytes as determinants of disease progression in inherited amyotrophic lateral sclerosis. *Nature Neurosci.* **11**, 251–253 (2008).
- Boillee, S. *et al.* Onset and progression in inherited ALS determined by motor neurons and microglia. *Science* **312**, 1389–1392 (2006).  
**This report showed that the expression of mutant SOD1 in microglia accelerates the progression of ALS, establishing a role for non-cell autonomous events in motor neuron degeneration in the disease.**
- Beers, D. R. *et al.* Wild-type microglia extend survival in PU.1 knockout mice with familial amyotrophic lateral sclerosis. *Proc. Natl Acad. Sci. USA* **103**, 16021–16026 (2006).
- Frakes, A. E. *et al.* Microglia induce motor neuron death via the classical NF- $\kappa$ B pathway in amyotrophic lateral sclerosis. *Neuron* **81**, 1009–1023 (2014).
- Harraz, M. M. *et al.* SOD1 mutations disrupt redox-sensitive Rac regulation of NADPH oxidase in a familial ALS model. *J. Clin. Invest.* **118**, 659–670 (2008).
- O'Rourke, J. G. *et al.* C9orf72 is required for proper macrophage and microglial function in mice. *Science* **351**, 1324–1329 (2016).



24. Jiang, J. *et al.* Gain of toxicity from ALS/FTD-linked repeat expansions in C9ORF72 is alleviated by antisense oligonucleotides targeting GGGGCC-containing RNAs. *Neuron* **90**, 535–550 (2016).  
**This report documented the amelioration of cognitive deficits in mice that express expanded hexanucleotide repeats following the intraventricular infusion of antisense oligonucleotides that reduce the levels of GGGGCC RNA transcripts and DPR proteins.**
25. Burberry, A. *et al.* Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. *Sci. Transl. Med.* **8**, 347ra93 (2016).
26. Kang, S. H. *et al.* Degeneration and impaired regeneration of gray matter oligodendrocytes in amyotrophic lateral sclerosis. *Nature Neurosci.* **16**, 571–579 (2013).
27. Lee, Y. *et al.* Oligodendroglia metabolically support axons and contribute to neurodegeneration. *Nature* **487**, 443–448 (2012).
28. Wang, L., Gutmann, D. H. & Roos, R. P. Astrocyte loss of mutant SOD1 delays ALS disease onset and progression in G85R transgenic mice. *Hum. Mol. Genet.* **20**, 286–293 (2011).
29. Howland, D. S. *et al.* Focal loss of the glutamate transporter EAAT2 in a transgenic rat model of SOD1 mutant-mediated amyotrophic lateral sclerosis (ALS). *Proc. Natl Acad. Sci. USA* **99**, 1604–1609 (2002).
30. Rothstein, J. D., Van Kammen, M., Levey, A. I., Martin, L. J. & Kuncel, R. W. Selective loss of glial glutamate transporter GLT-1 in amyotrophic lateral sclerosis. *Ann. Neurol.* **38**, 73–84 (1995).
31. Van Damme, P. *et al.* Astrocytes regulate GluR2 expression in motor neurons and their vulnerability to excitotoxicity. *Proc. Natl Acad. Sci. USA* **104**, 14825–14830 (2007).
32. Di Giorgio, F. P., Carrasco, M. A., Siao, M. C., Maniatis, T. & Eggan, K. Non-cell autonomous effect of glia on motor neurons in an embryonic stem cell-based ALS model. *Nature Neurosci.* **10**, 608–614 (2007).
33. Marchetto, M. C. *et al.* Non-cell-autonomous effect of human SOD1<sup>G37R</sup> astrocytes on motor neurons derived from human embryonic stem cells. *Cell Stem Cell* **3**, 649–657 (2008).
34. Nagai, M. *et al.* Astrocytes expressing ALS-linked mutated SOD1 release factors selectively toxic to motor neurons. *Nature Neurosci.* **10**, 615–622 (2007).
35. Haidet-Phillips, A. M. *et al.* Astrocytes from familial and sporadic ALS patients are toxic to motor neurons. *Nature Biotechnol.* **29**, 824–828 (2011).  
**This report demonstrated that motor neuron viability is compromised by astrocytes derived from post-mortem tissue in both familial ALS and sporadic ALS.**
36. Re, D. B. *et al.* Necroptosis drives motor neuron death in models of both sporadic and familial ALS. *Neuron* **81**, 1001–1008 (2014).
37. Lepore, A. C. *et al.* Focal transplantation-based astrocyte replacement is neuroprotective in a model of motor neuron disease. *Nature Neurosci.* **11**, 1294–1301 (2008).
38. Nishitoh, H. *et al.* ALS-linked mutant SOD1 induces ER stress- and ASK1-dependent motor neuron death by targeting Derlin-1. *Genes Dev.* **22**, 1451–1464 (2008).
39. Saxena, S., Cabuy, E. & Caroni, P. A role for motoneuron subtype-selective ER stress in disease manifestations of FALS mice. *Nature Neurosci.* **12**, 627–636 (2009).
40. Deng, H. X. *et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477**, 211–215 (2011).
41. Fecto, F. *et al.* SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Arch. Neurol.* **68**, 1440–1446 (2011).
42. Maruyama, H. *et al.* Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* **465**, 223–226 (2010).
43. Wild, P. *et al.* Phosphorylation of the autophagy receptor optineurin restricts *Salmonella* growth. *Science* **333**, 228–233 (2011).
44. Johnson, J. O. *et al.* Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* **68**, 857–864 (2010).
45. Parkinson, N. *et al.* ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). *Neurology* **67**, 1074–1077 (2006).
46. Chen, H. J. *et al.* Characterization of the properties of a novel mutation in VAPB in familial amyotrophic lateral sclerosis. *J. Biol. Chem.* **285**, 40266–40281 (2010).
47. Kabashi, E., Agar, J. N., Taylor, D. M., Minotti, S. & Durham, H. D. Focal dysfunction of the proteasome: a pathogenic factor in a mouse model of amyotrophic lateral sclerosis. *J. Neurochem.* **89**, 1325–1335 (2004).
48. Urushitani, M., Kurisu, J., Tsukita, K. & Takahashi, R. Proteasomal inhibition by misfolded mutant superoxide dismutase 1 induces selective motor neuron death in familial amyotrophic lateral sclerosis. *J. Neurochem.* **83**, 1030–1042 (2002).
49. Williamson, T. L. & Cleveland, D. W. Slowing of axonal transport is a very early event in the toxicity of ALS-linked SOD1 mutants to motor neurons. *Nature Neurosci.* **2**, 50–56 (1999).
50. Chen, X. J. *et al.* Proprioceptive sensory neuropathy in mice with a mutation in the cytoplasmic dynein heavy chain 1 gene. *J. Neurosci.* **27**, 14515–14524 (2007).
51. Persson, E. *et al.* A switch in retrograde signaling from survival to stress in rapid-onset neurodegeneration. *J. Neurosci.* **29**, 9903–9917 (2009).
52. Puls, I. *et al.* Mutant dynactin in motor neuron disease. *Nature Genet.* **33**, 455–456 (2003).
53. Gill, S. R. *et al.* Dynactin, a conserved, ubiquitously expressed component of an activator of vesicle motility mediated by cytoplasmic dynein. *J. Cell Biol.* **115**, 1639–1650 (1991).
54. Sutton, M. A. & Schuman, E. M. Dendritic protein synthesis, synaptic plasticity, and memory. *Cell* **127**, 49–58 (2006).
55. Alami, N. H. *et al.* Axonal transport of TDP-43 mRNA granules is impaired by ALS-causing mutations. *Neuron* **81**, 536–543 (2014).
56. Jucker, M. & Walker, L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45–51 (2013).
57. Grad, L. I. *et al.* Intermolecular transmission of superoxide dismutase 1 misfolding in living cells. *Proc. Natl Acad. Sci. USA* **108**, 16398–16403 (2011).
58. Münch, C., O'Brien, J. & Bertolotti, A. Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells. *Proc. Natl Acad. Sci. USA* **108**, 3548–3553 (2011).
59. Deng, H. X. *et al.* Conversion to the amyotrophic lateral sclerosis phenotype is associated with intermolecular linked insoluble aggregates of SOD1 in mitochondria. *Proc. Natl Acad. Sci. USA* **103**, 7142–7147 (2006).
60. Ayers, J. I., Fromholt, S. E., O'Neal, V. M., Diamond, J. H. & Borchelt, D. R. Prion-like propagation of mutant SOD1 misfolding and motor neuron disease spread along neuroanatomical pathways. *Acta Neuropathol.* **131**, 103–114 (2016).
61. Ravits, J. M. & La Spada, A. R. ALS motor phenotype heterogeneity, locality, and spread: deconstructing motor neuron degeneration. *Neurology* **73**, 805–811 (2009).
62. Urushitani, M. *et al.* Chromogranin-mediated secretion of mutant superoxide dismutase proteins linked to amyotrophic lateral sclerosis. *Nature Neurosci.* **9**, 108–118 (2006).
63. Lee, J. G., Takahama, S., Zhang, G., Tomarev, S. I. & Ye, Y. Unconventional secretion of misfolded proteins promotes adaptation to proteasome dysfunction in mammalian cells. *Nature Cell Biol.* **18**, 765–776 (2016).
64. Salajegheh, M. *et al.* Sarcoplasmic redistribution of nuclear TDP-43 in inclusion body myositis. *Muscle Nerve* **40**, 19–31 (2009).
65. Lippa, C. F. *et al.* Transactive response DNA-binding protein 43 burden in familial Alzheimer disease and Down syndrome. *Arch. Neurol.* **66**, 1483–1488 (2009).
66. Chanson, J. B. *et al.* TDP43-positive intraneuronal inclusions in a patient with motor neuron disease and Parkinson's disease. *Neurodegener. Dis.* **7**, 260–264 (2010).
67. Sreedharan, J. *et al.* TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* **319**, 1668–1672 (2008).  
**The first report of a germline TDP-43 mutation as a cause of familial ALS.**
68. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323**, 1208–1211 (2009).  
**This paper and ref. 71 were the first to identify mutations in the gene FUS that cause familial ALS.**
69. Kwiatkowski, T. J. Jr *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323**, 1205–1208 (2009).
70. Kim, H. J. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467–473 (2013).  
**The first report of a germline mutation in HNRNPA1 as a cause of familial ALS, and also the first demonstration that disease-causing mutations in low-complexity sequences can drive the hyperassembly of a membraneless organelle.**
71. Liu, Q. *et al.* Whole-exome sequencing identifies a missense mutation in hnRNPA1 in a family with flail arm ALS. *Neurology* <http://dx.doi.org/10.1212/WNL.0000000000003256> (2016).
72. Tollervey, J. R. *et al.* Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nature Neurosci.* **14**, 452–458 (2011).
73. Polymenidou, M. *et al.* Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nature Neurosci.* **14**, 459–468 (2011).
74. Lagier-Tourenne, C. *et al.* Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nature Neurosci.* **15**, 1488–1497 (2012).
75. Huelga, S. C. *et al.* Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep.* **1**, 167–178 (2012).
76. Johnson, J. O. *et al.* Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nature Neurosci.* **17**, 664–666 (2014).
77. Renton, A. E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257–268 (2011).
78. Hyman, A. A., Weber, C. A. & Julicher, F. Liquid–liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
79. Molliex, A. *et al.* Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).  
**This study demonstrated that low-complexity domains undergo phase separation that promotes the assembly of membraneless organelles such as stress granules and drives pathological fibrillization.**
80. Patel, A. *et al.* A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066–1077 (2015).  
**Ref. 81 and this paper were the first to show that pathological fibrils arise from liquid-like assemblies of proteins.**
81. Lin, Y., Protter, D. S., Rosen, M. K. & Parker, R. Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell* **60**, 208–219 (2015).
82. Lagier-Tourenne, C., Polymenidou, M. & Cleveland, D. W. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum. Mol. Genet.* **19**, R46–R64 (2010).
83. Shang, Y. & Huang, E. J. Mechanisms of FUS mutations in familial amyotrophic lateral sclerosis. *Brain Res.* **1647**, 65–78 (2016).
84. Murakami, T. *et al.* ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. *Neuron* **88**, 678–690 (2015).
85. Lim, S. M. *et al.* Directly converted patient-specific induced neurons mirror the neuropathology of FUS with disrupted nuclear localization in amyotrophic lateral sclerosis. *Mol. Neurodegener.* **11**, 8 (2016).
86. Ramaswami, M., Taylor, J. P. & Parker, R. Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell* **154**, 727–736 (2013).
87. Buratti, E. & Baralle, F. E. The molecular links between TDP-43 dysfunction and neurodegeneration. *Adv. Genet.* **66**, 1–34 (2009).

88. Ling, J. P., Pletnikova, O., Troncoso, J. C. & Wong, P. C. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science* **349**, 650–655 (2015).
89. Zhang, Z. *et al.* Downregulation of microRNA-9 in iPSC-derived neurons of FTD/ALS patients with TDP-43 mutations. *PLoS ONE* **8**, e76055 (2013).
90. Valdez, G., Heyer, M. P., Feng, G. & Sanes, J. R. The role of muscle microRNAs in repairing the neuromuscular junction. *PLoS ONE* **9**, e93140 (2014).
91. Williams, A. H. *et al.* MicroRNA-206 delays ALS progression and promotes regeneration of neuromuscular synapses in mice. *Science* **326**, 1549–1554 (2009).
92. Emde, A. *et al.* Dysregulated miRNA biogenesis downstream of cellular stress and ALS-causing mutations: a new mechanism for ALS. *EMBO J.* **34**, 2633–2651 (2015).
93. Cloutier, F., Marrero, A., O'Connell, C. & Morin, P., Jr. MicroRNAs as potential circulating biomarkers for amyotrophic lateral sclerosis. *J. Mol. Neurosci.* **56**, 102–112 (2015).
94. Morita, M. *et al.* A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology* **66**, 839–844 (2006).
95. Vance, C. *et al.* Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2–21.3. *Brain* **129**, 868–876 (2006).
96. Shatunov, A. *et al.* Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet Neurol.* **9**, 986–994 (2010).
97. Laakkovirta, H. *et al.* Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet Neurol.* **9**, 978–985 (2010).
98. Van Deerlin, V. M. *et al.* Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nature Genet.* **42**, 234–239 (2010).
99. van Es, M. A. *et al.* Genome-wide association study identifies 19p13.3 (*UNC13A*) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nature Genet.* **41**, 1083–1087 (2009).
100. Waite, A. J. *et al.* Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9orf72 hexanucleotide repeat expansion. *Neurobiol. Aging* **35**, 1779.e5–1779.e13 (2014).
101. Webster, C. P. *et al.* The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J.* **35**, 1656–1676 (2016).
102. Sellier, C. *et al.* Loss of C9orf72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J.* **35**, 1276–1297 (2016).
103. Lagier-Tourenne, C. *et al.* Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl Acad. Sci. USA* **110**, E4530–E4539 (2013).
104. Koppers, M. *et al.* C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann. Neurol.* **78**, 426–438 (2015).
105. Atanasio, A. *et al.* C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci. Rep.* **6**, 23204 (2016).
106. Chew, J. *et al.* C9orf72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science* **348**, 1151–1154 (2015).
107. Gendron, T. F. *et al.* Antisense transcripts of the expanded C9orf72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol.* **126**, 829–844 (2013).
108. Mori, K. *et al.* Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* **126**, 881–893 (2013).
109. Zu, T. *et al.* RAN proteins and RNA foci from antisense transcripts in C9orf72 ALS and frontotemporal dementia. *Proc. Natl Acad. Sci. USA* **110**, E4968–E4977 (2013).
110. La Spada, A. R. & Taylor, J. P. Repeat expansion disease: progress and puzzles in disease pathogenesis. *Nature Rev. Genet.* **11**, 247–258 (2010).
111. Haeusler, A. R. *et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**, 195–200 (2014).
112. Lee, Y. B. *et al.* Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* **5**, 1178–1186 (2013).
113. Mackenzie, I. R. *et al.* Dipeptide repeat protein pathology in C9orf72 mutation cases: clinico-pathological correlations. *Acta Neuropathol.* **126**, 859–879 (2013).
114. Xu, Z. *et al.* Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc. Natl Acad. Sci. USA* **110**, 7778–7783 (2013).
115. Cooper-Knock, J. *et al.* Antisense RNA foci in the motor neurons of C9orf72-ALS patients are associated with TDP-43 proteinopathy. *Acta Neuropathol.* **130**, 63–75 (2015).
116. Zhang, K. *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56–61 (2015).
117. Taylor, J. P. Neurodegenerative diseases: G-quadruplex poses quadruple threat. *Nature* **507**, 175–177 (2014).
118. Burguete, A. S. *et al.* GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. *eLife* **4**, e08881 (2015).
119. Ishiguro, A., Kimura, N., Watanabe, Y., Watanabe, S. & Ishihama, A. TDP-43 binds and transports G-quadruplex-containing mRNAs into neurites for local translation. *Genes Cells* **21**, 466–481 (2016).
120. Zu, T. *et al.* Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl Acad. Sci. USA* **108**, 260–265 (2011).
121. Ash, P. E. *et al.* Unconventional translation of C9orf72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77**, 639–646 (2013).  
**This paper and ref. 121 identified the DPR-protein pathology in C9 ALS-FTD tissue.**
122. Mori, K. *et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTD/ALS. *Science* **339**, 1335–1338 (2013).
123. Mackenzie, I. R. *et al.* Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9orf72 mutation carriers. *Acta Neuropathol.* **130**, 845–861 (2015).
124. Schludi, M. H. *et al.* Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. *Acta Neuropathol.* **130**, 537–555 (2015).
125. Kwon, I. *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139–1145 (2014).
126. Mizielinska, S. *et al.* C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* **345**, 1192–1194 (2014).
127. Wen, X. *et al.* Antisense proline-arginine RAN dipeptides linked to C9orf72-ALS/FTD form toxic nuclear aggregates that initiate *in vitro* and *in vivo* neuronal death. *Neuron* **84**, 1213–1225 (2014).
128. Freibaum, B. D. *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129–133 (2015).
129. Lee, K.-H. *et al.* C9orf72 dipeptide repeats impair the assembly, dynamics and function of membrane-less organelles. *Cell* **167**, 774–788 (2016).
130. Lin, Y. *et al.* Toxic PR poly-dipeptides encoded by the C9orf72 repeat expansion target LC domain polymers. *Cell* **167**, 789–802 (2016).
131. Yamakawa, M. *et al.* Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS. *Hum. Mol. Genet.* **24**, 1630–1645 (2015).
132. Zhang, Y. J. *et al.* Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol.* **128**, 505–524 (2014).
133. Zhang, Y. J. *et al.* C9orf72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nature Neurosci.* **19**, 668–677 (2016).
134. Jović, A. *et al.* Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nature Neurosci.* **18**, 1226–1229 (2015).
135. Donnelly, C. J. *et al.* RNA toxicity from the ALS/FTD C9orf72 expansion is mitigated by antisense intervention. *Neuron* **80**, 415–428 (2013).
136. Sareen, D. *et al.* Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9orf72 repeat expansion. *Sci. Transl. Med.* **5**, 208ra149 (2013).
137. O'Rourke, J. G. *et al.* C9orf72 BAC transgenic mice display typical pathologic features of ALS/FTD. *Neuron* **88**, 892–901 (2015).
138. Peters, O. M. *et al.* Human C9orf72 hexanucleotide expansion reproduces RNA foci and dipeptide repeat proteins but not neurodegeneration in BAC transgenic mice. *Neuron* **88**, 902–909 (2015).
139. Liu, Y. *et al.* C9orf72 BAC mouse model with motor deficits and neurodegenerative features of ALS/FTD. *Neuron* **90**, 521–534 (2016).
140. Brown, R. H., Jr. & Al-Chalabi, A. Endogenous retroviruses in ALS: a reawakening? *Sci. Transl. Med.* **7**, 307fs40 (2015).
141. Greenway, M. J. *et al.* ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. *Nature Genet.* **38**, 411–413 (2006).
142. Kabashi, E. *et al.* TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nature Genet.* **40**, 572–574 (2008).
143. Teyssou, E. *et al.* Mutations in *SQSTM1* encoding p62 in amyotrophic lateral sclerosis: genetics and neuropathology. *Acta Neuropathol.* **125**, 511–522 (2013).
144. Wu, C. H. *et al.* Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature* **488**, 499–503 (2012).
145. Smith, B. N. *et al.* Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS. *Neuron* **84**, 324–331 (2014).
146. Bannwarth, S. *et al.* A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain* **137**, 2329–2345 (2014).
147. Freischmidt, A. *et al.* Haploinsufficiency of *TBK1* causes familial ALS and frontotemporal dementia. *Nature Neurosci.* **18**, 631–636 (2015).

**Acknowledgements** We apologize for the many studies we were unable to cite because of space limitations. The authors gratefully acknowledge the artwork and editorial assistance provided by N. Nedelsky and H.-J. Kim and the provision of histopathology images by J. Ravits and S. Saberi. J.P.T. receives support from the Howard Hughes Medical Institute, the US National Institute of Neurological Disorders and Stroke (NINDS), the American Lebanese Syrian Associated Charities, Target ALS and the US ALS Association. R.H.B. receives support from NINDS, the ALS Association, ALS Finding a Cure, ALS ONE, the Angel Fund for ALS Research and Project ALS. D.W.C. receives salary support from the Ludwig Institute for Cancer Research and is supported by funding from NINDS (R01 NS27036), the ALS Association and Target ALS.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: see [go.nature.com/2exlcwy](http://go.nature.com/2exlcwy). Readers are welcome to comment on the online version of this paper at [go.nature.com/2exlcwy](http://go.nature.com/2exlcwy). Correspondence should be addressed to J.P.T. ([jpaul.taylor@stjude.org](mailto:jpaul.taylor@stjude.org)), R.H.B. ([robert.brown@umassmed.edu](mailto:robert.brown@umassmed.edu)) or D.W.C. ([dcleveland@ucsd.edu](mailto:dcleveland@ucsd.edu)).

**Reviewer Information** Nature thanks I. Dikic, J. Rothstein and the other anonymous reviewer(s) for their contribution to the peer review of this work.



# Defects in trafficking bridge Parkinson's disease pathology and genetics

Asa Abeliovich<sup>1,2,3</sup> & Aaron D. Gitler<sup>4</sup>

**Parkinson's disease is a debilitating, age-associated movement disorder. A central aspect of the pathophysiology of Parkinson's disease is the progressive demise of midbrain dopamine neurons and their axonal projections, but the underlying causes of this loss are unclear. Advances in genetics and experimental model systems have illuminated an important role for defects in intracellular transport pathways to lysosomes. The accumulation of altered proteins and damaged mitochondria, particularly at axon terminals, ultimately might overwhelm the capacity of intracellular disposal mechanisms. Cell-extrinsic mechanisms, including inflammation and prion-like spreading, are proposed to have both protective and deleterious functions in Parkinson's disease.**

Parkinson's disease (PD) is the most common neurodegenerative movement disorder, affecting about 1% of people aged 65 or older worldwide<sup>1</sup>. Clinical manifestations of PD include motor deficits such as rigidity, slowness in movement (bradykinesia), postural instability and a characteristic tremor at rest<sup>2</sup>. The motor symptoms of PD result from the selective loss of dopaminergic neurons in the pars compacta of the substantia nigra (SN) in the midbrain, as well as their axon terminals, which project to the dorsal striatum<sup>3</sup>. Dopamine replacement therapies, such as the dopamine precursor levodopa, typically lead to the relief of symptoms but lose their potency as the disease progresses.

A neuropathological hallmark of PD is the presence of intraneuronal proteinaceous inclusions, termed Lewy bodies (LBs) or Lewy neurites. These structures are enriched in filamentous forms of the synaptic protein  $\alpha$ -synuclein<sup>4,5</sup> as well as the small regulatory protein ubiquitin. Ubiquitin monomers or polyubiquitin chains are required for the proper intracellular trafficking and disposal of many proteins in cells, and the accumulation of ubiquitin is therefore consistent with defects in trafficking and disposal pathways<sup>6</sup>. Although LBs in the SN are a defining feature of PD in postmortem analysis, typically, these structures are also found more broadly in the PD brain<sup>7</sup>. The location of LBs seems to follow a specific pattern of progression: tissue from people with only mild clinical symptoms is typified by the presence of LBs mainly in lower regions of the brainstem such as the dorsal motor nucleus of the vagus nerve; by contrast, in people whose symptoms become increasingly severe, LBs are found in more rostral brain regions, including the SN, and ultimately the forebrain regions<sup>7</sup>. The systematic progression of LB pathology has been used as evidence to support the hypothesis that prion-like mechanisms of spreading might contribute to PD.

## Genetic causes of PD

Although PD was historically considered to be a sporadic disorder of unknown aetiology, it has become clear in the past two decades that a considerable proportion of cases (about 5–10%) are caused by familial genetic mutations<sup>8</sup>. Initial genetic studies focused on rare, inherited forms of PD in large families and mapped the causative mutations<sup>9–14</sup>. Exome and whole-genome sequencing approaches that use next-generation sequencing methods have enabled a further wave of rare, inherited causative mutations to be identified.

Genome-wide association studies (GWAS), in which genetic variants are queried across the entire human genome in large cohorts of affected

and unaffected people, have been transformative. For example, a large-scale meta-analysis of more than 19,000 people with PD and 100,000 controls that queried the role of about 8 million common genetic single nucleotide polymorphisms (SNPs) across the human genome identified 24 loci that are linked to an altered risk of developing the disease<sup>15</sup>. Although the effect size of each variant is typically modest (less than a 30% alteration in risk), together the variants have considerable impact<sup>16</sup>.

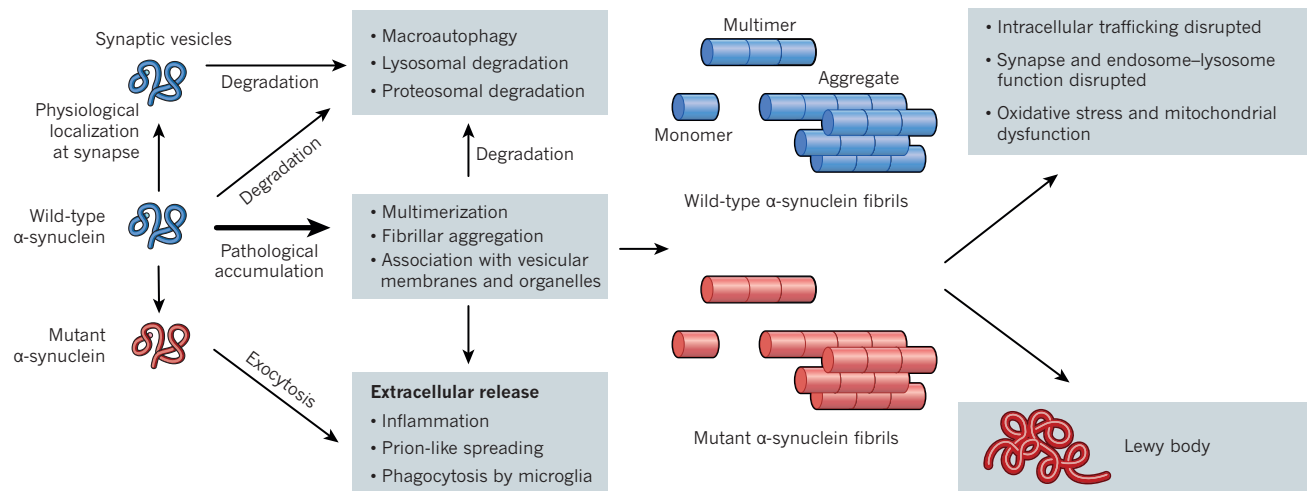
Genetic discoveries are beginning to illuminate cellular pathways and functions that are involved in the development of PD<sup>15</sup>, and the impairment of intracellular trafficking is emerging as a mechanistic link between many PD-associated genes. Here, we review the converging evidence that points to defects in the endosomal trafficking machinery, particularly at neuronal synapses, in the disruption of trafficking to the lysosome and in lysosomal dysfunction as pathological processes that contribute to PD, and we describe the putative role of  $\alpha$ -synuclein protein in these processes. We also relate these defects in intracellular trafficking to new data suggesting that extracellular mechanisms, including inflammation and the spread of disease-related proteins between cells, play important parts in the development of PD.

## $\alpha$ -Synuclein and intracellular trafficking

The first gene to be linked to familial PD was SNCA, which encodes the protein  $\alpha$ -synuclein<sup>9</sup>. Missense mutations in this gene, which alter the amino-acid sequence of the resulting 140-residue protein, cause rare, autosomal dominant inherited forms of PD-related diseases<sup>9</sup>. Similarly, genetic duplication or triplication of the SNCA locus can lead to familial, autosomal dominant forms of PD<sup>17,18</sup>. In a striking convergence of genetics and pathology, aggregates of  $\alpha$ -synuclein were discovered to be the building blocks of LBs, the pathological hallmark of PD<sup>5,19</sup>.

As well as the identification of missense mutations in  $\alpha$ -synuclein in rare, familial forms of PD, common genetic variants (typically defined as being found in more than 1% of a population) at the  $\alpha$ -synuclein locus have been associated with an increased lifetime risk of PD. For instance, in large population-based gene-association studies of people with PD and unaffected individuals, some genetic variations at the SNCA locus, including SNPs, were found to be enriched in those with PD, although modestly, which suggests that these variants increase the likelihood of developing the disease in the general population<sup>15,16,20</sup>. Such disease-associated SNPs do not seem to alter protein-coding sequences and their proposed mode of action is the modification of gene expression<sup>17,21,22</sup>.

<sup>1</sup>Department of Pathology and Cell Biology, Columbia University, New York, New York 10032, USA. <sup>2</sup>Department of Neurology, Columbia University, New York, New York 10032, USA. <sup>3</sup>Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, New York 10032, USA. <sup>4</sup>Department of Genetics, Stanford University, Stanford, California 94305, USA.



**Figure 1 | The proposed physiological and PD-associated pathological functions of  $\alpha$ -synuclein in neurons.** When generated at a moderate level, wild-type  $\alpha$ -synuclein (blue) associates with synaptic vesicles at axon terminals. Typically,  $\alpha$ -synuclein undergoes degradation either through lysosomal or proteasomal-led degradation. The pathological accumulation of wild-type  $\alpha$ -synuclein can result from an increase in its production — through high-risk, common genetic variants at the  $\alpha$ -synuclein gene that modify gene expression or through rare, familial gene multiplications — or from ineffective degradation or trafficking

An important question concerns how the accumulation of  $\alpha$ -synuclein contributes to PD pathogenesis. In other words, why is  $\alpha$ -synuclein toxic to cells? Studies in model systems that range from yeast cells to transgenic mice have shown that elevated levels of  $\alpha$ -synuclein disrupt numerous essential intracellular trafficking steps, including those at the endoplasmic reticulum<sup>23</sup>, the early and late endosomes<sup>24</sup> and the lysosome<sup>25,26</sup> (Fig. 1 and Box 1). The increased accumulation of  $\alpha$ -synuclein in long-term cultures of human induced pluripotent stem (iPS) cell-derived dopamine neurons, owing to genetic duplication or viral overexpression, leads to insufficient protein degradation as a consequence of the defective trafficking of important enzymes to lysosomes<sup>27</sup>. Upregulation of Ras-related protein Rab-1A, a prototypical regulator of vesicle trafficking (Box 1), is sufficient to rescue these trafficking deficits, which is reminiscent of the ability of a yeast homologue of Rab-1A to suppress  $\alpha$ -synuclein toxicity in yeast<sup>23</sup>. Elevated levels of  $\alpha$ -synuclein in neurons might therefore disrupt protein trafficking through vesicles or the endosome and, in particular, to the lysosome for degradation<sup>28</sup>.  $\alpha$ -Synuclein has also been implicated in a separate intracellular pathway for trafficking proteins to the lysosome, termed macroautophagy (Box 2).

The actions of  $\alpha$ -synuclein on vesicular transport processes might reflect its biophysical properties.  $\alpha$ -Synuclein in solution is mostly a disordered monomer with a propensity to bind lipid membranes<sup>29,30</sup>, which could underlie its role in the disruption of vesicular trafficking, as well as its physiological functions in regulating the release of synaptic vesicles. In certain contexts,  $\alpha$ -synuclein can also assume a variety of soluble or insoluble oligomeric forms<sup>30–32</sup>, although the precise physiological or pathological roles of such structures are contentious. Insoluble  $\alpha$ -synuclein aggregates and fibrils with  $\beta$ -sheet secondary structure are a main constituent of LBs and seem to be neurotoxic<sup>33,34</sup>.

In model systems such as neurons derived from human embryonic stem cells, the presence of PD-causing mutations in  $\alpha$ -synuclein has been associated with the accumulation of reactive oxygen and nitrogen species<sup>25,35</sup>. This could reflect the build-up of defective mitochondria, which are essential for the production of ATP, the energy currency of the cell, but also produce reactive species as a by-product. Mitochondria are disposed of typically through a process termed mitophagy, in which these organelles are encompassed by double-membraned structures, forming autophagosomes that later fuse to lysosomes for degradation<sup>36,37</sup>.

pathways. These defective mechanisms have also been associated with rare, familial PD-associated mutations in  $\alpha$ -synuclein (red) as well as mutations in trafficking-associated genes linked to PD and environmental factors such as ageing (not shown). An excess of  $\alpha$ -synuclein in the form of monomers, multimers or aggregates can disrupt intracellular trafficking and synaptic function and contributes to the formation of LBs. Furthermore, the extracellular release of  $\alpha$ -synuclein through exocytosis might lead to inflammation and the spread of PD-associated lesions such as LBs through prion-like mechanisms.

Overexpression of  $\alpha$ -synuclein in dopamine neurons of the mouse mid-brain has been reported to impede mitophagy<sup>38</sup>. Several genes involved in familial PD, including *PARK2* (encoding the protein parkin), *PINK1*, *FBXO7* and *PARK7* (also known as *DJ-1*), have important roles in the execution of mitophagy (Fig. 2), which are reviewed elsewhere<sup>39</sup>.

$\alpha$ -Synuclein is released from cells — particularly when accumulated in excess — through various pathways, including those dedicated to secretion and those involved in the exocytosis of lysosomes<sup>40–43</sup>. Such extracellular release of  $\alpha$ -synuclein could serve as a backup disposal mechanism to complement intracellular lysosomes. Neighbouring cells such as innate-immune microglia are thought to have an enhanced capacity for phagocytosis and lysosomal degradation. But released  $\alpha$ -synuclein might also have detrimental roles by inducing inflammation directly<sup>44–46</sup>, by acting as a chemoattractant that draws inflammatory microglia towards damaged neurons<sup>47</sup> or by helping to spread PD pathology.

Many studies have sought to identify the physiological functions of  $\alpha$ -synuclein in neurons and to relate these to the pathology of PD. For example, knockout mice that are deficient in  $\alpha$ -synuclein show potentiated vesicular transmitter release<sup>48</sup>. *In vitro* studies suggest that  $\alpha$ -synuclein might normally function as a chaperone to modulate the activity of the synaptic-vesicle fusion machinery<sup>49,50</sup>. And the affinity of  $\alpha$ -synuclein for phospholipids, which are enriched on vesicle membranes<sup>51,52</sup>, probably facilitates its role as a regulator of vesicle trafficking.

In a broader sense, perhaps the toxic and normal functions of  $\alpha$ -synuclein are intimately linked. At physiological levels,  $\alpha$ -synuclein might function as a regulator of vesicle fusion and neurotransmitter release at the synapse (Fig. 3). However, its accumulation beyond a certain threshold (for example, in people with SNPs that elevate the level of  $\alpha$ -synuclein) could lead to the inappropriate deployment of this regulatory function at the synapse or the promiscuous inhibition of other steps in vesicular trafficking<sup>53</sup> (Fig. 3), including those in the endosome-lysosome pathway.

### **LRRK2, GBA and lysosomal trafficking**

Autosomal dominant, inherited mutations in the gene *LRRK2* are the most prevalent genetic lesions that underlie familial forms of PD<sup>13,14</sup>. *LRRK2* encodes a large, multidomain protein that contains kinase, GTPase and other protein-interaction domains. The most common



## BOX 1

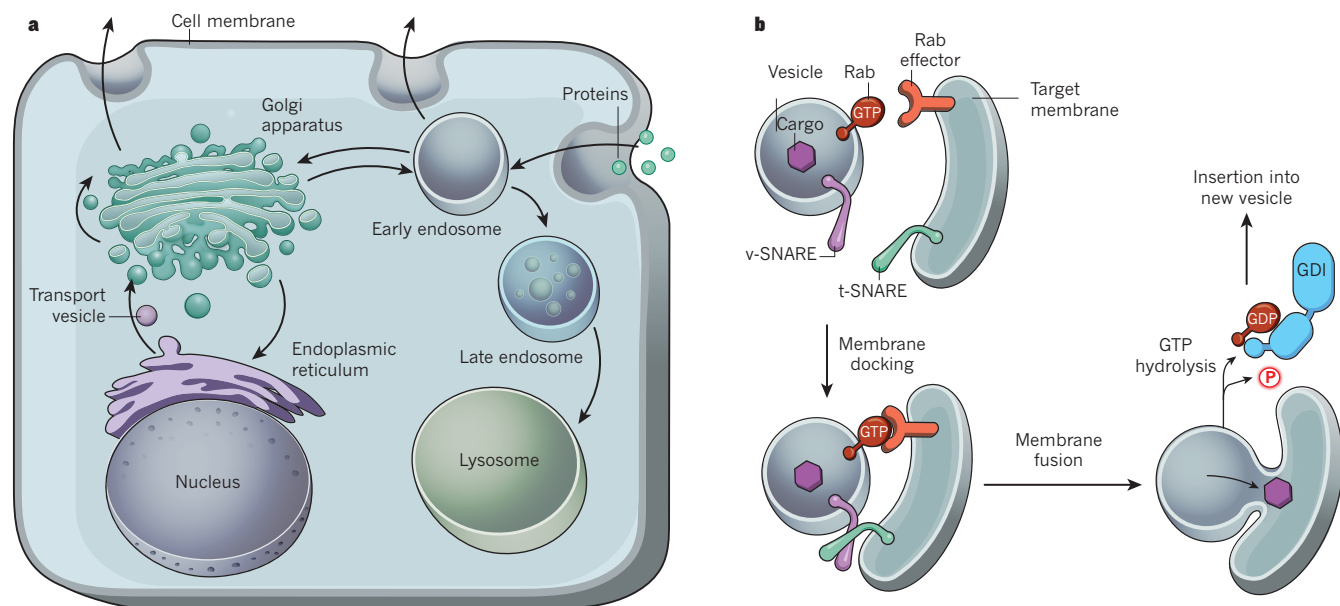
# Overview of intracellular trafficking

● **Secretory and endocytic pathways** Select protein cargoes are transported from their site of production, the endoplasmic reticulum, to their final destinations — either the cell-surface membrane or other cellular organelles — through the secretory pathway (Box Fig. a). The cargoes are packaged into transport vesicles, which bud from the membrane and are delivered to the target compartment, where they fuse. As proteins are transported through the Golgi apparatus, some will be modified by the addition of carbohydrates, which provide further information for targeting proteins to different parts of the cell.

● **Endocytosis** Proteins and other molecules are transported into the cell through endocytosis (Box Fig. a). During this process, part of the plasma membrane becomes invaginated and pinches away from the membrane to form a vesicle that incorporates the proteins. Invagination can be enabled by a coating composed of the protein clathrin, which is then removed. The vesicles are delivered to other membrane-bound vesicles called endosomes, where they are sorted. Some proteins are returned to the plasma membrane or sent to the Golgi apparatus, a process that is regulated by the retromer complex, which includes VPS35. Proteins that are destined for degradation are delivered to the late endosome (or a specialized type of endosome known as a multivesicular body) that subsequently fuses with the

lysosome. The contents of the endosome are released to be degraded by enzymes in the acidic environment of the lysosome. Lysosomal contents can also be released into the extracellular environment through secretion.

● **Vesicle docking and fusion with target membranes** Small proteins called Rab GTPases are involved in targeting vesicles to the correct membranes<sup>135</sup> (Box Fig. b). A Rab protein on the surface of the vesicle interacts with an adaptor protein (Rab effector) that is located on the target membrane. This interaction causes the vesicle to dock with the appropriate target membrane. Specificity is achieved by using different types of Rab proteins and effectors for each type of vesicle and at each step of the secretory pathway; for instance, RAB1A functions selectively in transport between the endoplasmic reticulum and the Golgi apparatus. After the vesicle has docked with the target membrane, transmembrane proteins called SNAP receptors (SNAREs) interact, which brings the two membranes together to enable fusion. On vesicle fusion, GTP is hydrolysed to GDP, and the Rab protein is extracted from the membrane by the protein Rab GDP dissociation inhibitor (GDI) and reinserted into the membrane of a new vesicle so that the cycle can be repeated. SNAP, soluble NSF attachment protein; t-SNARE, target SNARE; v-SNARE, vesicle SNARE.



mutation associated with *LRRK2* in Western countries is G2019S, which occurs in the kinase domain and is thought to disinhibit the kinase activity — towards at least a subset of substrates<sup>54,55</sup>. The expression of mutant *LRRK2* in both cells and animal models leads to defective endosome-to-lysosome trafficking, the accumulation of abnormal lysosomal structures and a reduction in the number of neurite processes in neurons<sup>55–57</sup>. In addition to defective vesicular trafficking to the lysosome, mutant *LRRK2* has been implicated in other pathological processes such as dysregulated protein translation<sup>58</sup>.

In support of its role in vesicular trafficking, *LRRK2* has been shown to interact with several proteins from the Rab family<sup>56,57,59,60</sup>, which are important regulators of vesicular intracellular trafficking (Box 1). Among these Rab proteins are Ras-related protein Rab-7L1 (also known as RAB29) (refs 56 and 61), encoded by the *PARK16* locus and therefore associated with the risk of developing PD, and the structurally related protein RAB32, which has been linked to another neurodegenerative

disease, amyotrophic lateral sclerosis<sup>59</sup>. These two Rab proteins are part of a subfamily that is implicated in trafficking to the lysosomes and to lysosome-like organelles. Phosphoprotein analyses have also implicated *LRRK2* in phosphorylating and modulating other Rab proteins, including RAB3A, RAB8A, RAB10 and RAB12 (ref. 60). Mutant *LRRK2* that causes PD might harbour a potentiated kinase activity towards these Rabs, probably leading to altered interactions with downstream effectors of Rab and its regulatory proteins, as well as to perturbations in vesicular transport at several steps (Box 1). Future studies will be required to define whether and how some of these substrates for *LRRK2* kinase activity contribute to PD.

Neither transgenic mice that express PD-associated mutant *LRRK2* nor *LRRK2*-knockout mice consistently show a robust PD-like phenotype, such as the progressive loss of dopamine neurons. However, *LRRK2*-knockout mice do show the features of age-associated lysosomal pathology — particularly in proximal tubule cells of the kidney and type II pneumocytes of the lung<sup>62–64</sup>. These data reinforce the link between the

## BOX 2

# Cellular components of macromolecular degradation

● **Lysosome** An acidic organelle that harbours a variety of enzymes for the degradation of proteins, lipids, sugars and other macromolecules. These are trafficked from the external environment (through mechanisms such as endocytosis)<sup>79</sup>, from endocytic compartments or from the cytoplasm (by macroautophagy). The fusion of lysosomes with late endosomes or autophagosomes is a process that requires the SNARE complex and RAB7 (as well as the homotypic fusion and protein sorting (HOPS) complex for autophagy)<sup>136</sup> that seems to be impaired by the accumulation of  $\alpha$ -synuclein<sup>28,130</sup>.

● **Macroautophagy** An evolutionarily conserved process by which certain constituents of the cytoplasm such as defective proteins, aggregates or organelles are engulfed by double-membrane-bound phagophores, which envelop cargoes to form autophagosomes that can fuse to lysosomes<sup>137</sup>. The initiation of autophagy is negatively regulated by the mTOR signalling pathway and inhibitors of this pathway have therefore been considered as potential therapeutic

drugs. However, the accumulation of  $\alpha$ -synuclein and other mechanisms associated with the development of PD might impair the fusion of autophagosomes with lysosomes, and increased initiation could therefore be harmful.

● **Mitophagy** The macroautophagic degradation of defective mitochondria. On the depolarization of mitochondria, PINK1 becomes stabilized and able to recruit the ubiquitin ligase parkin<sup>39</sup> as well as autophagy receptors such as optineurin<sup>132,138</sup>.

● **Proteasome** A multiprotein complex that degrades short-lived cytoplasmic proteins typically marked by polyubiquitin chains<sup>6</sup>.

● **Ubiquitin** A short polypeptide that can be attached to cellular proteins by ligases. Ubiquitin can be assembled into polyubiquitin chains on target proteins that serve as postal codes for various destinations, including the proteasome or lysosome, that are determined partly by the molecular nature of the inter-ubiquitin linkages<sup>6</sup>.

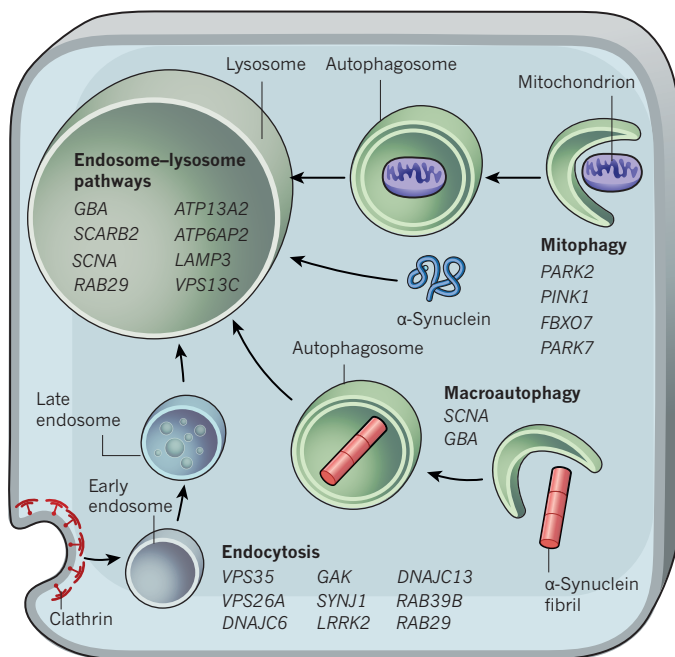
function of LRRK2 and lysosome-associated trafficking. Because people with mutations in *LRRK2* eventually accumulate  $\alpha$ -synuclein in their brains, it has been proposed that  $\alpha$ -synuclein represents a cargo that is relevant for LRRK2-associated trafficking. Alternatively, the effect of *LRRK2* mutations on the accumulation of  $\alpha$ -synuclein might be indirect, the result of more general endosome and lysosome dysfunction. Although many studies support a model in which PD-associated mutations in *LRRK2* lead to the disinhibition of kinase activity, and therefore a gain of function, it is possible that PD is caused by a reduction in kinase activity; for instance, kinase activity towards a subset of substrates might

be reduced in the context of disinhibition, or the non-kinase activities of LRRK2 could be modified instead. The gain-of-function versus loss-of-function issue is even more vexing with respect to the impact of common genetic variants associated with the risk of developing PD, such as those found at the *LRRK2* locus, because typically these variants do not change the coding sequence of the gene and might only alter expression subtly.

Autosomal recessive mutations in the gene *GBA*, which encodes a lysosomal hydrolase called glucocerebrosidase, lead to the defective breakdown of the glycosphingolipid glucoceramide (also known as glucosylceramide (GlcCer)) into ceramide and glucose, and cause Gaucher's disease, a lysosomal storage disorder with neurological features that include parkinsonism as well as a variety of other clinical findings<sup>65</sup>. Heterozygous carriers of *GBA* mutations have a considerably higher risk of developing PD (about threefold to eightfold greater)<sup>66,67</sup>. A deficiency in glucocerebrosidase can lead to neurodegeneration, either indirectly through general lysosome dysfunction and the failure of endosome-lysosome or autophagosome-lysosome fusion<sup>68</sup> or through a more direct link between the accumulation of GlcCer and  $\alpha$ -synuclein. For instance, GlcCer has been shown to stabilize  $\alpha$ -synuclein oligomers. Glucocerebrosidase deficiency and the consequent accumulation of GlcCer might also damage cells through the over-activation of the endoplasmic-reticulum-associated degradation (ERAD) pathway and the disruption of other cellular homeostatic mechanisms such as stress-associated calcium release<sup>68</sup>. And a reduction in glucocerebrosidase activity might disrupt the generation of sphingolipids in the central nervous system (CNS) through the ceramide synthesis pathway, leading to further membrane-associated dyshomeostasis<sup>26,68</sup>.

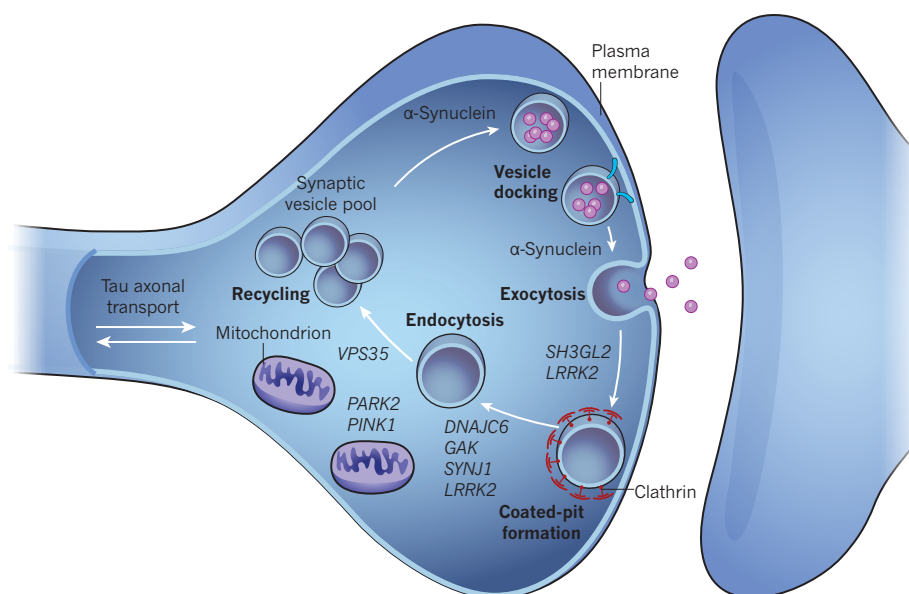
Human iPS-cell-derived neuronal models of PD that are associated with a mutation in *GBA* show an increased accumulation of both  $\alpha$ -synuclein<sup>26,68</sup> and GlcCer. Furthermore, neuronal cells that overexpress  $\alpha$ -synuclein have a defect in the trafficking of glucocerebrosidase to the lysosome<sup>26,27</sup>, which points to a positive feedback loop. Notably, a cofactor required for the import of glucocerebrosidase into the lysosome, lysosome membrane protein 2 (LIMP-2), which is encoded by the gene *SCARB2*, has also been linked by GWAS to the risk of developing PD<sup>69</sup>. A deficiency in LIMP-2 leads to the defective transport of glucocerebrosidase, as well as lysosomal dysfunction and  $\alpha$ -synuclein accumulation<sup>15,70</sup>.

Other PD-related genes identified by familial genetic studies or GWAS also support roles for altered trafficking to the lysosome and defective lysosome integrity in the pathogenesis of PD (Fig. 2). Familial autosomal mutations in the gene *ATP13A2*, which encodes a transmembrane P-type ATPase that is localized to lysosomes and late endosomes and enriched in the brain, underlie an early onset form of parkinsonism with dementia



**Figure 2 | PD-related genes associated with trafficking to the lysosome.** Genes that encode intracellular trafficking components are associated with common sporadic and familial forms of PD, as well as related syndromes that share some of the clinical features of PD. Most of these genes are known to affect trafficking to the lysosome in the context of late endosome-to-lysosome pathways, clathrin-dependent endocytosis, macroautophagy or mitophagy. Wild-type  $\alpha$ -synuclein (blue) can also enter lysosomes through chaperone-mediated autophagy<sup>118</sup>.





**Figure 3 | The synaptic vesicle cycle is implicated in PD.** Synaptic vesicles typically cycle through several stages: docking at the plasma membrane of the synaptic terminal; exocytosis through fusion with the plasma membrane in a calcium-dependent manner; the formation of a coated pit at the plasma membrane, which is enabled by a protein coating that is composed of numerous clathrin proteins and other protein components; endocytosis; recycling into a readily releasable pool of synaptic vesicles; and, re-docking with the plasma membrane<sup>133</sup>. In vesicle cycling, mitochondria are essential for generating ATP, buffering calcium and enabling other functions at the synaptic terminal, and proper axonal trafficking must also be maintained. Many PD-related genes are implicated in these processes (red).  $\alpha$ -Synuclein usually localizes to synaptic vesicles and might modulate the pool of vesicles and vesicle docking with the membrane. In PD,  $\alpha$ -synuclein accumulates inappropriately and too robustly interferes with the priming of synaptic vesicles, which leads to a decrease in the size of the pool of releasable vesicles.

and other symptoms<sup>71</sup>. Mutations in *ATP13A2* lead to a reduction in the function of lysosomes and to the accumulation of abnormal lysosomes; they also sensitize cells to oxidative stress, mitochondrial dysfunction and manganese toxicity<sup>72</sup>. In cell models, mutations in this gene have been reported to cause the accumulation of  $\alpha$ -synuclein, however, such accumulation was not seen in animal models that were deficient in *ATP13A2* (ref. 73). *ATP13A2* might interact with synaptotagmin-11 (encoded by the gene *SYT11*), which has been implicated in lysosomal function and exocytosis<sup>74,75</sup> and associated by GWAS with the risk of developing PD<sup>16</sup>. Mutations in *ATP6AP2*, an essential accessory component of the vacuolar-type  $H^+$  ATPase that is required for lysosome acidification and function, have been associated with an X-linked form of parkinsonism with spasticity<sup>76</sup>. In animal models, depletion of the *ATP6AP2* protein led to neurodegeneration with evidence of defective protein degradation through the autophagy–lysosome pathway<sup>77</sup>. Rare, autosomal recessive loss-of-function mutations in *VPS13C*, a gene encoding a component of the vesicular machinery that was specifically implicated in sorting to the lysosome on the basis of the analysis of a yeast orthologue, have been associated with familial PD<sup>78</sup>, and common variants at this locus have also been associated with a risk of developing sporadic PD<sup>15</sup>.

Such studies, driven by genetic associations with PD and related syndromes, point to altered trafficking to lysosomes and altered lysosomal function as potent mechanisms that underlie the loss of neurons in PD.

### Endocytosis and retromer dysfunction in PD

The endocytosis of membrane proteins proceeds stepwise from the cell-surface plasma membrane, through clathrin-coat-mediated or clathrin-independent routes<sup>79</sup>, to early endosome and late-endosome vesicles and eventually leads either to the recycling of cargo — such as through the retromer pathway to the Golgi apparatus — or to the degradation of cargo in the lysosome (Box 1). Mutations in several components of the endosome–lysosome trafficking machinery have been associated with familial forms of PD. The discovery of autosomal dominant mutations in the gene *VPS35*, which encodes a component of the retromer complex, has implicated retromer dysfunction in familial PD<sup>80,81</sup>. The retromer protein complex is required for the recycling of certain cargo proteins in early endosomes to the *trans*-Golgi network or plasma membrane. The cargo-selection function of the retromer complex is orchestrated by a trimer that is comprised of protein subunits VPS35, VPS26A or VPS26B and VPS29, and a dimer of sorting nexin proteins, including SNX1 and SNX2, which bind and curve endosomal membranes, presumably to enable the sequestration of selected cargo<sup>82,83</sup>.

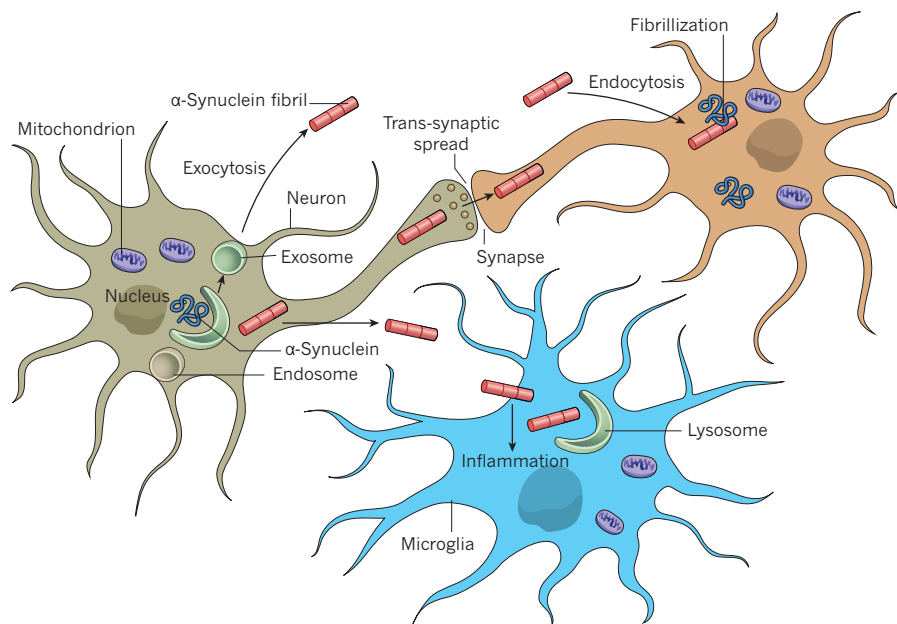
At a cellular level, PD-associated mutations in *VPS35* have been

linked to defects in vesicular trafficking and neuronal toxicity<sup>56,84–86</sup>. Several retromer cargo proteins, such as the cation-independent mannose-6-phosphate receptor, are essential for the delivery of the main component enzymes of lysosomes, and retromer dysfunction therefore leads to the disruption of lysosomal trafficking and integrity<sup>56,87</sup>, which is reminiscent of the findings seen in the context of mutations in the genes *LRRK2* and *GBA*.

When part of the multimeric retromer complex, VPS35 might function as a scaffold that binds to the complex and links it to other trafficking machinery<sup>88</sup>, and PD-associated mutations in *VPS35* possibly disrupt these interactions. For instance, the retromer complex associates with the multicomponent WASH complex, and together these enable the sequestration of cargo in the early endosomes and its sorting to appropriate destinations. PD-associated mutations in *VPS35* have been shown to disrupt interactions of the WASH complex<sup>84,89</sup>, leading to proteotoxic stress<sup>89</sup>.

VPS35 and retromer dysfunction have also been connected directly to the pathological effects of  $\alpha$ -synuclein, as a loss of VPS35 function can sensitize cells to the accumulation of  $\alpha$ -synuclein by interfering with the degradation machinery in a range of model systems, including yeast and transgenic mice<sup>85,90,91</sup>. Potentiated retromer function might suppress the altered trafficking and toxicity that is associated with mutations in *LRRK2* (ref. 56) or the overexpression of  $\alpha$ -synuclein<sup>85</sup>, which suggests a potential therapeutic avenue.

Remarkably, at least five further genes that are associated with the endocytic pathway have been linked to PD. Three functionally related proteins — DnaJ homolog subfamily C member 13 (also known as RME-8), encoded by the gene *DNAJC13* (ref. 92), putative tyrosine-protein phosphatase auxilin (auxilin 1), encoded by *DNAJC6* (ref. 93), and cyclin-G-associated kinase (auxilin 2), encoded by *GAK*<sup>15</sup> — play integral parts in clathrin-mediated endocytosis at the plasma membrane by helping to uncoat clathrin, including at the synaptic terminal (Fig. 3). And the genes that encode these proteins have either been associated with familial PD (probably through loss-of-function mutations in *DNAJC13* (ref. 92) and *DNAJC6* (ref. 93) or with the risk of developing sporadic PD (through common variants in *GAK*<sup>15</sup>). VPS35 and RME-8 have both been shown to interact with the WASH complex<sup>84,89</sup>, which further links the PD-associated genes that encode them to a common pathway. The gene *SYNJ1* (refs 94 and 95) encodes the protein synaptotagmin-1, which is required for the shedding of clathrin adaptors in the endosomal pathway, and mutations in *SYNJ1* are linked to defects in late endosomal and lysosomal trafficking and to a familial syndrome with Parkinson's disease and seizures<sup>94,95</sup>. Loss-of-function mutations in the gene *RAB39B*, which encodes a Rab-family GTPase that localizes to early endosomes, have



**Figure 4 | Extracellular  $\alpha$ -synuclein and the prion hypothesis.**  $\alpha$ -Synuclein can be released into the extracellular environment from neurons, especially in the context of its excessive accumulation, which can result from defective lysosomal function. Monomers, multimers or fibrillar insoluble aggregates of  $\alpha$ -synuclein can be released through several routes, including lysosomal exocytosis, trans-synaptic spreading or the death of neurons (not shown). Extracellular fibrillar  $\alpha$ -synuclein can then gain entry to the cytoplasm of distant neurons by directly penetrating the plasma membrane, through bulk endocytosis<sup>134</sup> or by other means, where it nucleates the fibrillization of other native  $\alpha$ -synuclein molecules that are present throughout the cells. Extracellular  $\alpha$ -synuclein might also induce inflammation through the activation of Toll-like receptors found on the surface of innate immune cells such as microglia. Innate immune cells are capable of clearing extracellular  $\alpha$ -synuclein through phagocytosis and lysosomal degradation.

been described in familial forms of parkinsonism with cognitive impairment and  $\alpha$ -synuclein pathology<sup>96</sup>. Furthermore, RAB39B was specifically implicated in the accumulation of  $\alpha$ -synuclein in neurons. Together, these human-genetics findings point to defective endocytic trafficking and sorting, culminating in lysosomal dysfunction and defects in proteostasis, as convergent mechanisms that underlie PD or clinically related syndromes.

Important questions remain regarding the nature of the cargo with impaired transport that underlies PD. What is the normal fate of this cargo? Why does its missorting lead to lysosomal dysfunction and toxicity? And where does such cargo accumulate when trafficking is impaired? In some contexts, such as with mutations in *SCARB2* or *VPS35*, the delivery of essential lysosomal components is proposed to be selectively disrupted. Alternatively, missorted proteins might accumulate in toxic forms or cause a more general disruption in trafficking.

### Synaptic trafficking is at the epicentre of PD

Postmortem studies of the brain support the idea that the generation of defects at axon processes and terminals is an early event in PD, and many PD-associated genes are implicated in synaptic function and trafficking at axon terminals (Fig. 3).  $\alpha$ -Synuclein is localized mostly to pre-synaptic terminals in the normal brain and it has a regulatory function in synaptic vesicle release<sup>50,53,97</sup>. *LRK2* has been specifically implicated in the presynaptic regulation of endophilin A1, which facilitates endocytosis through clathrin uncoating at the synaptic terminal<sup>98</sup>. *VPS35* and *LRK2* both seem to be necessary for proper presynaptic function and have been implicated in the trafficking of regulatory components at the synaptic terminal<sup>99</sup>. It is conceivable that these proteins also modulate  $\alpha$ -synuclein activity or localization in some contexts. Mutations in the gene that encodes TMEM230, a transmembrane protein enriched at synaptic and recycling vesicles in neurons, have been described in autosomal dominant familial PD<sup>100</sup>.

As well as synaptic terminal dysfunction, findings from human genetics and functional studies<sup>28</sup> have also implicated defective axonal transport to and from synaptic terminals in PD (Fig. 3). Several GWAS have strongly linked the risk of developing PD to common variants at the *MATP* locus, which encodes the microtubule-associated protein tau<sup>15,101,102</sup>. Although aggregates of tau are associated historically with Alzheimer's disease and frontotemporal dementia, the accumulation of tau has also been reported in LBs and in brains affected by PD<sup>103</sup>.

### Trafficking, spreading and inflammation

Aside from impairments in intracellular transport, transcellular mechanisms, including inflammation and prion-like spreading, have emerged

as potential drivers of PD. These cell-extrinsic mechanisms might serve to propagate intracellular trafficking defects.

### Prion-like spreading of $\alpha$ -synuclein pathology

A concept that has the potential to transform research into neurodegenerative diseases is the hypothesis that proteins linked to neurodegeneration such as  $\alpha$ -synuclein and tau might undergo prion-like spreading<sup>104,105</sup>. Prions are well established as the protein-based infectious agents that underlie the transmissible spongiform encephalopathies (for example, bovine spongiform encephalopathy in cattle and Creutzfeldt-Jakob disease in humans). In these rare but devastating diseases, the major prion protein (known as PrP) converts from the normal soluble form to an aggregated, self-templating infectious form. This process initiates an inexorable spread of aggregates and contingent neurodegeneration throughout the brain. Accumulating evidence suggests that this phenomenon might extend to more common neurodegenerative diseases such as PD and Alzheimer's disease.

Several studies have focused on the mechanisms of  $\alpha$ -synuclein spreading using *in vitro* and *in vivo* model systems (Fig. 4). Notably, the spread of  $\alpha$ -synuclein-containing aggregates to endogenous brain neurons seems to require the expression of intracellular  $\alpha$ -synuclein: spreading was not observed in  $\alpha$ -synuclein-knockout mice<sup>33</sup>, which is consistent with a model in which exogenously injected fibrils spread by templating the aggregation of endogenous  $\alpha$ -synuclein in a prion-like manner. Notably, in parallel to these  $\alpha$ -synuclein-based models, studies focusing on other neurodegenerative proteins such as tau also point to pathological spreading as a mechanism for disease progression<sup>106</sup>.

Key questions remain about the precise mechanisms by which  $\alpha$ -synuclein aggregates are released from and enter neurons *in vivo*. Dysfunction in intracellular vesicular trafficking that is associated with PD might also affect the spread of  $\alpha$ -synuclein between cells. Emerging data support this idea: for instance, mutations in *GBA* seem to promote the exocytosis of  $\alpha$ -synuclein<sup>107</sup>, and mutations in *ATP13A2* affect the spread of  $\alpha$ -synuclein through exosome biogenesis<sup>42,43</sup> whereas those in *VPS35* might promote the endocytosis of aggregated  $\alpha$ -synuclein<sup>85</sup>. It is important to emphasize that the *in vivo* model systems in which neuron-to-neuron spread of PD pathological features (for example,  $\alpha$ -synuclein aggregates) have been reported are inherently non-physiological because they are based on the exogenous injection of  $\alpha$ -synuclein protein. Whether such spreading actually occurs in humans and, if so, whether it has a causative role in PD, remains unresolved<sup>108</sup>.



### Inflammation is a double-edged sword in PD

Inflammation has also been implicated as a mechanism of spreading  $\alpha$ -synuclein pathology. Although inflammatory activation in the CNS is classically induced in the context of infections, it has also been associated with neurodegenerative disorders such as PD<sup>109</sup>. As a first line of defence, innate immune cells in the CNS, including resident microglia and infiltrating peripheral macrophages, express pattern-recognition receptors such as Toll-like receptors that recognize evolutionarily conserved pathogen-associated molecular patterns (PAMPs)<sup>110</sup>. As well as pathogenic species, innate immune cells can also detect endogenous danger-associated molecular patterns (DAMPs), such as cytoplasmic or lysosomal contents that have been released into the extracellular space from injured cells or through exocytotic mechanisms<sup>111</sup>.

Signs of inflammation have long been noted in postmortem examinations in the context of PD; such features include morphologically altered microglia with fewer ramified processes, the induction of certain cell-surface markers and the release of inflammatory cytokines and chemokines, including interleukin-1 $\beta$ , tumour-necrosis factor- $\alpha$

and interferon- $\gamma$ <sup>109</sup>. However, there is debate about the extent to which inflammation serves protective or disease-causing roles in PD. Innate immune cells can actively clear debris, including protein aggregates such as  $\alpha$ -synuclein, in the context of models of PD such as glucosylceramidase insufficiency<sup>107,112</sup>. It has been suggested that  $\alpha$ -synuclein might function as a DAMP in the context of PD, leading to increased inflammation<sup>113</sup> (Fig. 4). And if such inflammation causes the further release of  $\alpha$ -synuclein, the process could perpetuate itself.

Innate immune cells are professional phagocytic cells with a high capacity for internalizing and degrading cell debris and protein aggregates through the lysosome pathway<sup>45,114</sup>. It is therefore reasonable to postulate that at the point in PD when neurons become over-burdened with proteotoxic stress, the lysosomal degradation machinery in adjoining innate immune cells can have an important supportive role<sup>115</sup>. However, after the lysosomal degradation machinery of even the innate immune cells has been overwhelmed by cell debris such as  $\alpha$ -synuclein aggregates, microglia and other innate immune cells probably become over-activated and cause damage to neurons<sup>109,116</sup>. Almost all of the trafficking-related genes

**Table 1 | Intracellular trafficking genes that contribute to PD**

Gene	Familial or risk	Genetic inheritance	Protein	Proposed function of encoded protein	Trafficking steps	References	Therapeutic approaches
<i>SNCA</i>	Both	AD	$\alpha$ -Synuclein	Potential SNARE-complex assembly chaperone	Endosome-lysosome, synaptic	49, 50	Anti- $\alpha$ -synuclein antibody <sup>131</sup> , $\alpha$ -synuclein-reducing small molecules, aggregation inhibitors
<i>LRRK2</i>	Both	AD	LRRK2	Kinase, GTPase	Endosome-lysosome, synaptic	60	LRRK2 kinase inhibitors <sup>63,129</sup>
<i>GBA</i>	Risk	NA	Glucocerebrosidase	Lysosomal glucocerebrosidase enzyme	Lysosome, ER, Golgi apparatus	26, 27	GBA enzyme activity enhancement <sup>27,128</sup>
<i>SCARB2</i>	Risk	NA	LIMP2	Chaperone for glucocerebrosidase trafficking	Lysosome	69	None known
<i>ATP13A2</i>	Familial	AR	ATPase 13A2	ATPase cation transporter	Lysosome	43, 71	None known
<i>ATP6AP2</i>	Familial	X	ATP6AP2	ATPase proton transporter	Lysosome	76	None known
<i>SYT11</i>	Risk	NA	Synaptotagmin-11	Transmembrane regulator of lysosome-autophagosome fusion and exocytosis	Lysosome	NA	None known
<i>VPS13C</i>	Both	AR	VPS13C	Endosomal sorting, mitophagy	Endosome-lysosome	15, 78	None known
<i>VPS35</i>	Familial	AD	VPS35	Retromer subunit	Endosome-lysosome, Golgi apparatus	84	Small-molecule retromer activity boosters <sup>130</sup>
<i>DNAJC13</i>	Familial	AD	RME-8	Co-chaperone function in clathrin uncoating during endosomal transport	Endocytosis, synaptic	92	None known
<i>DNAJC6</i>	Familial	AR	Auxilin-1	Co-chaperone function in clathrin uncoating during endosomal transport	Endocytosis, synaptic	93	None known
<i>GAK</i>	Risk	NA	Auxilin-2	Co-chaperone function in clathrin uncoating during endosomal transport	Endocytosis, synaptic	15	None known
<i>SYNJ1</i>	Familial	AR	Synaptojanin-1	Phosphoinositide phosphatase	Endocytosis, synaptic	94, 95	None known
<i>RAB39B</i>	Familial	X	RAB39B	Rab GTPase	Endosome	96	None known
<i>TMEM230</i>	Familial	AD	Transmembrane protein 230	Transmembrane secretory or recycling vesicle protein	Endosome-lysosome, synaptic	100	None known
<i>MAPT</i>	Risk	NA	Tau	Microtubule-associated protein	Axonal transport	15, 101, 102	None known
<i>LAMP3</i>	Risk	NA	Lysosome-associated membrane protein 3	Regulator of protein degradation during cellular unfolded protein response	Lysosome	16	None known
<i>PINK1</i>	Familial	AR	PINK1	Kinase, phosphorylates parkin and ubiquitin to regulate mitophagy	Mitochondria	37, 132	None known
<i>PARK2</i>	Familial	AR	Parkin	Ubiquitin ligase, regulator of mitophagy	Mitochondria	37, 132	None known
<i>RAB29</i>	Risk	NA	RAB7-L1	Rab GTPase	Endosome-lysosome, Golgi apparatus	56	None known

AD, autosomal dominant; AR, autosomal recessive; ER, endoplasmic reticulum; NA, not applicable; X, X linked.

associated with PD, including *GBA*, *LRRK2* and *VPS35*, have also been shown to regulate trafficking and vesicular function in innate immune cells. Therefore, disease-causing trafficking dysfunctions in PD might, in part, occur in innate immune cells, leading to proteostasis and inflammation that culminate in non-autonomous neurotoxicity.

### Linking trafficking defects to neuronal loss

A remarkable number of familial and sporadic PD genes are involved, directly or indirectly, in endosome–lysosome trafficking (Table 1). Trafficking-related genes are implicated not only in rare, familial forms of PD on the basis of the identification of inherited mutations, but also in non-familial cases of PD through common variants at these genes that modify the risk of developing PD; in fact, PD genes such as *SNCA* and *LRRK2* are linked both to familial PD and to the risk of developing common, sporadic disease.

The proclivity of CNS neurons, and more specifically of midbrain dopamine neurons and their axon terminals, to develop the pathological hallmarks of PD is surprising given that the endosome–lysosome trafficking mechanisms implicated by the genetics of PD genetics are likely to affect many types of cells, and that most known PD-associated genes (Table 1) are expressed broadly. Neurons are forced to preserve homeostasis with ageing through degradation mechanisms because they do not divide, unlike most other eukaryotic cell types. Furthermore, dopamine itself might induce oxidative and nitrosative damage through the production of reactive species<sup>35</sup>, which places dopaminergic axon terminals and their synaptic vesicles at considerable risk. Dopamine might interact directly with  $\alpha$ -synuclein at axon terminals<sup>117,118</sup> to mediate its toxic effects<sup>119</sup>.

The unique architecture of midbrain dopaminergic neurons could present a particular challenge for trafficking owing to the exuberance and vast number of their axons and synaptic terminals: for example, a single human dopaminergic neuron can harbour more than 1 million axon terminals. The onset of PD pathology therefore might be a consequence of ineffective proteostasis at axon terminals as the result of limited endosome–lysosome pathway functions. Historically, axon terminals were considered to be mostly devoid of lysosomes but several studies have challenged that view<sup>120,121</sup>. Interestingly, the familial-PD genes *PARK2* and *PINK1* play an essential part in the recruitment of the autophagic and lysosomal degradation machinery to defective mitochondria to initiate mitophagy at distal axons<sup>121</sup>. Dopaminergic neurons might also be affected by unique physiological stressors: for example, their reliance on particularly efficient types of calcium channels, including  $\text{Ca}_v1.3$  L-type channels, leads to increased calcium uptake, which probably places further strain on cellular compartments such as mitochondria and lysosomes<sup>122</sup>.

Notably, genes and pathways involved in trafficking have also been associated with other neurodegenerative disorders such as Alzheimer's disease, frontotemporal dementia and amyotrophic lateral sclerosis. This convergence points to common biological pathways in cells on which to focus future studies<sup>123–126</sup>. It is important to ask how genes that are implicated in a common pathway could affect seemingly distinct neuropathologies: for example, the retromer-complex genes *VPS35* and *SORL1* are associated with PD and AD, respectively. Distinct cargoes could be involved in the same pathway in different disorders, or particular cell types affected by the disorders might rely disproportionately on specific pathway-associated disease genes.

### Potential therapeutic targets

Defective trafficking to the lysosome is the most prominent mechanism of pathogenesis that links PD-associated genetic variants and mutations, and it therefore has high priority as a therapeutic target. However, diverse strategies are being pursued because there is a complex interplay between the various intracellular trafficking pathways. Lysosomal storage disorders, which typically are severe, childhood inherited syndromes caused by defective lysosome function, might be mechanistically related to PD, and PD could therefore represent a mild and late-onset *forme fruste* of

a lysosomal storage disorder. Given the overlap between PD and such disorders, some of the same therapeutic strategies might be applicable to both. Drugs that enhance lysosome function or trafficking, including  $\beta$ -cyclodextrins, or the transduction of genes such as *TFEB*, which encodes the transcription factor EB<sup>127</sup>, might provide therapeutic benefits in the context of PD. Strategies to selectively improve the function or delivery of glucosylceramidase, including functional activators of the enzyme<sup>27</sup> or molecular chaperones that stabilize its structure have been considered<sup>128</sup>, but their clinical efficacy has not been described. Because several of the most common pathogenic mutations in *LRRK2*, including G2019S, result in promiscuous LRRK2 kinase activity<sup>54</sup>, approaches to inhibit such activity are being pursued<sup>63,129</sup>. However, the relevance of altered LRRK2 kinase activity to the development of PD is unresolved and, paradoxically, inhibitors of LRRK2 kinase activity have been associated with lysosomal disorders *in vivo*<sup>63</sup>. The emerging role of *VPS35* mutations and retromer-complex function in PD means that approaches that boost retromer activity could be a promising therapeutic avenue<sup>130</sup>. Improved autophagic flux, for example, through inhibition of the serine–threonine-protein kinase mTOR with rapamycin, has been evaluated, but such a strategy could be counterproductive if the main defect is downstream in the pathway at the fusion of autophagosomes to lysosomes.

A particular focus has been placed on therapies that selectively affect  $\alpha$ -synuclein accumulation or inhibitors of its fibrillization, although the precise structure of the toxic form of  $\alpha$ -synuclein is under debate. A therapeutic small molecule called NAB2, which might broadly suppress defects in endosomal trafficking that occur downstream of an excess of  $\alpha$ -synuclein, has been described using a yeast-based screen<sup>25</sup>. Inhibitors of oxidative or nitrosative damage, which might promote toxic structures, have also been proposed<sup>35</sup>. Given the potential role of extracellular  $\alpha$ -synuclein in PD — as a promoter of prion-like spreading or inflammation — antibody-based therapeutic drugs for the clearance of  $\alpha$ -synuclein are being pursued extensively<sup>121</sup> (as they are for proteins involved in other neurodegenerative diseases), but their value is yet to be validated<sup>128</sup>. Therapeutic efforts that target  $\alpha$ -synuclein might be broadly effective if the protein is an essential downstream effector of PD-associated neuron loss, regardless of the initial insult. Alternatively, it is possible that mechanisms that are independent of  $\alpha$ -synuclein, and perhaps related to trafficking dysfunction, drive PD-associated neurotoxicity in most cases. ■

Received 11 May; accepted 2 September 2016.

1. Kalia, L. V. & Lang, A. E. Parkinson's disease. *Lancet* **386**, 896–912 (2015).
2. Goedert, M., Spillantini, M. G., Del Tredici, K. & Braak, H. 100 years of Lewy pathology. *Nature Rev. Neurol.* **9**, 13–24 (2013).
3. Hornykiewicz, O. Dopamine (3-hydroxytyramine) in the central nervous system and its relation to the Parkinson syndrome in man [in German]. *Dtsch. Med. Wochenschr.* **87**, 1807–1810 (1962).
4. Dickson, D. W. Parkinson's disease and parkinsonism: neuropathology. *Cold Spring Harb. Perspect. Med.* **2**, a009258 (2012).
5. Spillantini, M. G. *et al.*  $\alpha$ -Synuclein in Lewy bodies. *Nature* **388**, 839–840 (1997).
6. Ciechanover, A. Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nature Rev. Mol. Cell Biol.* **6**, 79–87 (2005).
7. Braak, H. & Braak, E. Pathoanatomy of Parkinson's disease. *J. Neurol.* **247** (suppl. 2), II3–II10 (2000).
8. Lin, M. K. & Farrer, M. J. Genetics and genomics of Parkinson's disease. *Genome Med.* **6**, 48 (2014).
9. Polymeropoulos, M. H. *et al.* Mapping of a gene for Parkinson's disease to chromosome 4q21–q23. *Science* **274**, 1197–1199 (1996).
10. Kitada, T. *et al.* Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* **392**, 605–608 (1998).
11. Bonifati, V. *et al.* Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science* **299**, 256–259 (2003).
12. Valente, E. M. *et al.* Hereditary early-onset Parkinson's disease caused by mutations in *PINK1*. *Science* **304**, 1158–1160 (2004).
13. Zimprich, A. *et al.* Mutations in *LRRK2* cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* **44**, 601–607 (2004).
14. Paisán-Ruiz, C. *et al.* Cloning of the gene containing mutations that cause *PARK8*-linked Parkinson's disease. *Neuron* **44**, 595–600 (2004).
15. Nalls, M. A. *et al.* Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nature Genet.* **46**, 989–993 (2014).
16. International Parkinson Disease Genomics Consortium. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* **377**, 641–649 (2011).



17. Singleton, A. B. *et al.*  $\alpha$ -Synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841 (2003).
18. Chartier-Harlin, M. C. *et al.*  $\alpha$ -Synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* **364**, 1167–1169 (2004).
19. Lee, V. M. & Trojanowski, J. Q. Mechanisms of Parkinson's disease linked to pathological  $\alpha$ -synuclein: new targets for drug discovery. *Neuron* **52**, 33–38 (2006).
20. Farrer, M. *et al.*  $\alpha$ -Synuclein gene haplotypes are associated with Parkinson's disease. *Hum. Mol. Genet.* **10**, 1847–1851 (2001).
21. Rhinn, H. *et al.* Alternative  $\alpha$ -synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology. *Nature Commun.* **3**, 1084 (2012).
22. Soldner, F. *et al.* Parkinson-associated risk variant in distal enhancer of  $\alpha$ -synuclein modulates target gene expression. *Nature* **533**, 95–99 (2016).
23. Cooper, A. A. *et al.*  $\alpha$ -Synuclein blocks ER–Golgi traffic and Rab1 rescues neuron loss in Parkinson's models. *Science* **313**, 324–328 (2006).
24. Outeiro, T. F. & Lindquist, S. Yeast cells provide insight into  $\alpha$ -synuclein biology and pathobiology. *Science* **302**, 1772–1775 (2003).
25. Chung, C. Y. *et al.* Identification and rescue of  $\alpha$ -synuclein toxicity in Parkinson patient-derived neurons. *Science* **342**, 983–987 (2013).
- Cortical neurons generated from people with a mutation in  $\alpha$ -synuclein exhibit cellular features that could be predicted from the results of an unbiased screen of a yeast  $\alpha$ -synuclein toxicity model and that could be reversed using a small-molecule 'hit' from the screen.**
26. Mazzulli, J. R. *et al.* Gaucher disease glucocerebrosidase and  $\alpha$ -synuclein form a bidirectional pathogenic loop in synucleinopathies. *Cell* **146**, 37–52 (2011).
- A cellular pathway links Gaucher's disease and PD through impaired lysosomal-enzyme targeting and protein degradation, which suggests that therapeutic targets and strategies might be effective for both disorders.**
27. Mazzulli, J. R., Zinke, F., Isacson, O., Studer, L. & Krainc, D.  $\alpha$ -Synuclein-induced lysosomal dysfunction occurs through disruptions in protein trafficking in human midbrain synucleinopathy models. *Proc. Natl Acad. Sci. USA* **113**, 1931–1936 (2016).
- Long-term cultures of human midbrain dopamine neurons from people with PD or Gaucher's disease exhibit defects in the trafficking of important lysosomal enzymes, which could be rescued by boosting the transport of vesicles.**
28. Volpicelli-Daley, L. A. *et al.* Formation of  $\alpha$ -synuclein Lewy neurite-like aggregates in axons impedes the transport of distinct endosomes. *Mol. Biol. Cell* **25**, 4010–4023 (2014).
29. Theillet, F. X. *et al.* Structural disorder of monomeric  $\alpha$ -synuclein persists in mammalian cells. *Nature* **530**, 45–50 (2016).
30. Burré, J., Sharma, M. & Sudhof, T. C.  $\alpha$ -Synuclein assembles into higher-order multimers upon membrane binding to promote SNARE complex formation. *Proc. Natl Acad. Sci. USA* **111**, E4274–E4283 (2014).
31. Bartels, T., Choi, J. G. & Selkoe, D. J.  $\alpha$ -Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature* **477**, 107–110 (2011).
32. Dettmer, U. *et al.* Parkinson-causing  $\alpha$ -synuclein missense mutations shift native tetramers to monomers as a mechanism for disease initiation. *Nature Commun.* **6**, 7314 (2015).
33. Luk, K. C. *et al.* Pathological  $\alpha$ -synuclein transmission initiates Parkinson-like neurodegeneration in nontransgenic mice. *Science* **338**, 949–953 (2012).
- A single, intrastriatal injection of  $\alpha$ -synuclein fibrils in wild-type mice initiates the templating and propagation of pathological aggregates, the loss of dopaminergic neurons and motor deficits, providing a direct mechanistic link between the spread of  $\alpha$ -synuclein and the pathogenesis of PD.**
34. Volpicelli-Daley, L. A. *et al.* Exogenous  $\alpha$ -synuclein fibrils induce Lewy body pathology leading to synaptic dysfunction and neuron death. *Neuron* **72**, 57–71 (2011).
35. Ryan, S. D. *et al.* Isogenic human iPSC Parkinson's model shows nitrosative stress-induced dysfunction in MEF2–PGC1 $\alpha$  transcription. *Cell* **155**, 1351–1364 (2013).
36. Pickrell, A. M. *et al.* Endogenous Parkin preserves dopaminergic substantia nigral neurons following mitochondrial DNA mutagenic stress. *Neuron* **87**, 371–381 (2015).
37. Youle, R. J. & Narendra, D. P. Mechanisms of mitophagy. *Nature Rev. Mol. Cell Biol.* **12**, 9–14 (2011).
38. Chen, L., Xie, Z., Turkson, S. & Zhuang, X. A53T human  $\alpha$ -synuclein overexpression in transgenic mice induces pervasive mitochondria macroautophagy defects preceding dopamine neuron degeneration. *J. Neurosci.* **35**, 890–905 (2015).
39. Pickrell, A. M. & Youle, R. J. The roles of PINK1, Parkin, and mitochondrial fidelity in Parkinson's disease. *Neuron* **85**, 257–273 (2015).
40. Liu, B. *et al.* The polarisome is required for segregation and retrograde transport of protein aggregates. *Cell* **140**, 257–267 (2010).
41. Lee, H. J. *et al.* Autophagic failure promotes the exocytosis and intercellular transfer of  $\alpha$ -synuclein. *Exp. Mol. Med.* **45**, e22 (2013).
42. Kong, S. M. *et al.* Parkinson's disease-linked human PARK9/ATP13A2 maintains zinc homeostasis and promotes  $\alpha$ -synuclein externalization via exosomes. *Hum. Mol. Genet.* **23**, 2816–2833 (2014).
43. Tsunemi, T., Hamada, K. & Krainc, D. ATP13A2/PARK9 regulates secretion of exosomes and  $\alpha$ -synuclein. *J. Neurosci.* **34**, 15281–15287 (2014).
44. Fellner, L. *et al.* Toll-like receptor 4 is required for  $\alpha$ -synuclein dependent activation of microglia and astroglia. *Glia* **61**, 349–360 (2013).
45. Rannikko, E. H., Weber, S. S. & Kahle, P. J. Exogenous  $\alpha$ -synuclein induces toll-like receptor 4 dependent inflammatory responses in astrocytes. *BMC Neurosci.* **16**, 57 (2015).
46. Kim, C. *et al.* Neuron-released oligomeric  $\alpha$ -synuclein is an endogenous agonist of TLR2 for paracrine activation of microglia. *Nature Commun.* **4**, 1562 (2013).
47. Wang, S. *et al.*  $\alpha$ -Synuclein, a chemoattractant, directs microglial migration via H<sub>2</sub>O<sub>2</sub>-dependent Lyn phosphorylation. *Proc. Natl Acad. Sci. USA* **112**, E1926–E1935 (2015).
48. Abeliovich, A. *et al.* Mice lacking  $\alpha$ -synuclein display functional deficits in the nigrostriatal dopamine system. *Neuron* **25**, 239–252 (2000).
49. Chandra, S., Gallardo, G., Fernandez-Chacon, R., Schluter, O. M. & Sudhof, T. C.  $\alpha$ -Synuclein cooperates with CSP $\alpha$  in preventing neurodegeneration. *Cell* **123**, 383–396 (2005).
50. Burré, J. *et al.*  $\alpha$ -Synuclein promotes SNARE-complex assembly *in vivo* and *in vitro*. *Science* **329**, 1663–1667 (2010).
- $\alpha$ -Synuclein is shown to function as a chaperone that facilitates the maintenance of SNARE complexes at the synapse, and *Snc* *Sncb* *Sncg* triple-knockout mice exhibited neurological impairments and deficits in SNARE-complex assembly, revealing an important physiological function for  $\alpha$ -synuclein of relevance to neurodegeneration.**
51. Kubo, S. *et al.* A combinatorial code for the interaction of  $\alpha$ -synuclein with membranes. *J. Biol. Chem.* **280**, 31664–31672 (2005).
52. Nakamura, K. *et al.* Direct membrane association drives mitochondrial fission by the Parkinson disease-associated protein  $\alpha$ -synuclein. *J. Biol. Chem.* **286**, 20710–20726 (2011).
53. Gitler, A. D. & Shorter, J. Prime time for  $\alpha$ -synuclein. *J. Neurosci.* **27**, 2433–2434 (2007).
54. West, A. B. *et al.* Parkinson's disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity. *Proc. Natl Acad. Sci. USA* **102**, 16842–16847 (2005).
55. MacLeod, D. *et al.* The familial Parkinsonism gene *LRRK2* regulates neurite process morphology. *Neuron* **52**, 587–593 (2006).
56. MacLeod, D. A. *et al.* *RAB7L1* interacts with *LRRK2* to modify intraneuronal protein sorting and Parkinson's disease risk. *Neuron* **77**, 425–439 (2013).
57. Dodson, M. W., Leung, L. K., Lone, M., Lizzio, M. A. & Guo, M. Novel ethyl methanesulfonate (EMS)-induced null alleles of the *Drosophila* homolog of *LRRK2* reveal a crucial role in endolysosomal functions and autophagy *in vivo*. *Dis. Model. Mech.* **7**, 1351–1363 (2014).
58. Martin, I. *et al.* Ribosomal protein s15 phosphorylation mediates *LRRK2* neurodegeneration in Parkinson's disease. *Cell* **157**, 472–485 (2014).
59. Waschbüsch, D. *et al.* *LRRK2* transport is regulated by its novel interacting partner Rab32. *PLoS ONE* **9**, e111632 (2014).
60. Steger, M. *et al.* Phosphoproteomics reveals that Parkinson's disease kinase *LRRK2* regulates a subset of Rab GTPases. *eLife* **5**, e12813 (2016).
- A proteomics approach identifies specific Rab GTPases as physiological targets of the kinase *LRRK2*; pathogenic variants of *LRRK2* increase the phosphorylation of such targets, providing a mechanism that connects mutations in *LRRK2* to defects in vesicle-trafficking steps that are orchestrated by Rab GTPases and their effectors.**
61. Beilina, A. *et al.* Unbiased screen for interactors of leucine-rich repeat kinase 2 supports a common pathway for sporadic and familial Parkinson disease. *Proc. Natl Acad. Sci. USA* **111**, 2626–2631 (2014).
62. Tong, Y. *et al.* Loss of leucine-rich repeat kinase 2 causes impairment of protein degradation pathways, accumulation of  $\alpha$ -synuclein, and apoptotic cell death in aged mice. *Proc. Natl Acad. Sci. USA* **107**, 9879–9884 (2010).
63. Fujii, R. N. *et al.* Effect of selective *LRRK2* kinase inhibition on nonhuman primate lung. *Sci. Transl. Med.* **7**, 273ra15 (2015).
64. Herzig, M. C. *et al.* *LRRK2* protein levels are determined by kinase function and are crucial for kidney and lung homeostasis in mice. *Hum. Mol. Genet.* **20**, 4209–4223 (2011).
65. Sidransky, E. Gaucher disease and parkinsonism. *Mol. Genet. Metab.* **84**, 302–304 (2005).
66. Zuckerman, S. *et al.* Carrier screening for Gaucher disease: lessons for low-penetrance, treatable diseases. *J. Am. Med. Assoc.* **298**, 1281–1290 (2007).
67. Goker-Alpan, O. *et al.* Parkinsonism among Gaucher disease carriers. *J. Med. Genet.* **41**, 937–940 (2004).
68. Schöndorf, D. C. *et al.* iPSC-derived neurons from GBA1-associated Parkinson's disease patients show autophagic defects and impaired calcium homeostasis. *Nature Commun.* **5**, 4028 (2014).
69. Do, C. B. *et al.* Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genet.* **7**, e1002141 (2011).
70. Rothaug, M. *et al.* LIMP-2 expression is critical for  $\beta$ -glucocerebrosidase activity and  $\alpha$ -synuclein clearance. *Proc. Natl Acad. Sci. USA* **111**, 15573–15578 (2014).
71. Ramirez, A. *et al.* Hereditary parkinsonism with dementia is caused by mutations in *ATP13A2*, encoding a lysosomal type 5 P-type ATPase. *Nature Genet.* **38**, 1184–1191 (2006).
72. Gitler, A. D. *et al.*  $\alpha$ -Synuclein is part of a diverse and highly conserved interaction network that includes *PARK9* and manganese toxicity. *Nature Genet.* **41**, 308–315 (2009).
73. Kett, L. R. *et al.*  $\alpha$ -Synuclein-independent histopathological and motor deficits in mice lacking the endolysosomal Parkinsonism protein *Atp13a2*. *J. Neurosci.* **35**, 5724–5742 (2015).
74. Usenovic, M., Tresse, E., Mazzulli, J. R., Taylor, J. P. & Krainc, D. Deficiency of *ATP13A2* leads to lysosomal dysfunction,  $\alpha$ -synuclein accumulation, and neurotoxicity. *J. Neurosci.* **32**, 4240–4246 (2012).
75. Bento, C. F., Ashkenazi, A., Jimenez-Sanchez, M. & Rubinsztein, D. C. The Parkinson's disease-associated genes *ATP13A2* and *SYT11* regulate autophagy via a common pathway. *Nature Commun.* **7**, 11803 (2016).
76. Korvatska, O. *et al.* Altered splicing of *ATP6AP2* causes X-linked parkinsonism with spasticity (XPDS). *Hum. Mol. Genet.* **22**, 3259–3268 (2013).
77. Dubos, A. *et al.* Conditional depletion of intellectual disability and Parkinsonism candidate gene *ATP6AP2* in fly and mouse induces cognitive impairment and neurodegeneration. *Hum. Mol. Genet.* **24**, 6736–6755 (2015).

78. Lesage, S. *et al.* Loss of VPS13C function in autosomal-recessive Parkinsonism causes mitochondrial dysfunction and increases PINK1/Parkin-dependent mitophagy. *Am. J. Hum. Genet.* **98**, 500–513 (2016).
  79. McMahon, H. T. & Boucrot, E. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nature Rev. Mol. Cell Biol.* **12**, 517–533 (2011).
  80. Zimprich, A. *et al.* A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. *Am. J. Hum. Genet.* **89**, 168–175 (2011).
  81. Vilarinho-Güell, C. *et al.* VPS35 mutations in Parkinson disease. *Am. J. Hum. Genet.* **89**, 162–167 (2011).
  82. Parks, W. T. *et al.* Sorting nexin 6, a novel SNX, interacts with the transforming growth factor- $\beta$  family of receptor serine-threonine kinases. *J. Biol. Chem.* **276**, 19332–19339 (2001).
  83. Seaman, M. N. & Freeman, C. L. Analysis of the retromer complex–WASH complex interaction illuminates new avenues to explore in Parkinson disease. *Commun. Integr. Biol.* **7**, e29483 (2014).
  84. Zavadsky, E. *et al.* Mutation in VPS35 associated with Parkinson's disease impairs WASH complex association and inhibits autophagy. *Nature Commun.* **5**, 3828 (2014).
  85. Dhungel, N. *et al.* Parkinson's disease genes VPS35 and EIF4G1 interact genetically and converge on  $\alpha$ -synuclein. *Neuron* **85**, 76–87 (2015).
  86. Tsika, E. *et al.* Parkinson's disease-linked mutations in VPS35 induce dopaminergic neurodegeneration. *Hum. Mol. Genet.* **23**, 4621–4638 (2014).
  87. Arighi, C. N., Hartnell, L. M., Aguilar, R. C., Haft, C. R. & Bonifacio, J. S. Role of the mammalian retromer in sorting of the cation-independent mannose 6-phosphate receptor. *J. Cell Biol.* **165**, 123–133 (2004).
  88. Hierro, A. *et al.* Functional architecture of the retromer cargo-recognition complex. *Nature* **449**, 1063–1067 (2007).
  89. McGough, I. J. *et al.* Retromer binding to FAM21 and the WASH complex is perturbed by the Parkinson disease-linked VPS35(D620N) mutation. *Curr. Biol.* **24**, 1670–1676 (2014).
  90. Miura, E. *et al.* VPS35 dysfunction impairs lysosomal degradation of  $\alpha$ -synuclein and exacerbates neurotoxicity in a *Drosophila* model of Parkinson's disease. *Neurobiol. Dis.* **71**, 1–13 (2014).
  91. Tang, F. L. *et al.* VPS35 in dopamine neurons is required for endosome-to-Golgi retrieval of Lamp2a, a receptor of chaperone-mediated autophagy that is critical for  $\alpha$ -synuclein degradation and prevention of pathogenesis of Parkinson's disease. *J. Neurosci.* **35**, 10613–10628 (2015).
  92. Vilarinho-Güell, C. *et al.* DNAJC13 mutations in Parkinson disease. *Hum. Mol. Genet.* **23**, 1794–1801 (2014).
  93. Edvardson, S. *et al.* A deleterious mutation in DNAJC6 encoding the neuronal-specific clathrin-uncoating co-chaperone auxilin, is associated with juvenile parkinsonism. *PLoS ONE* **7**, e36458 (2012).
  94. Quadri, M. *et al.* Mutation in the SYNJ1 gene associated with autosomal recessive, early-onset Parkinsonism. *Hum. Mutat.* **34**, 1208–1215 (2013).
  95. Krebs, C. E. *et al.* The Sac1 domain of SYNJ1 identified mutated in a family with early-onset progressive Parkinsonism with generalized seizures. *Hum. Mutat.* **34**, 1200–1207 (2013).
  96. Wilson, G. R. *et al.* Mutations in RAB39B cause X-linked intellectual disability and early-onset Parkinson disease with  $\alpha$ -synuclein pathology. *Am. J. Hum. Genet.* **95**, 729–735 (2014).
  97. Nemani, V. M. *et al.* Increased expression of  $\alpha$ -synuclein reduces neurotransmitter release by inhibiting synaptic vesicle recluster after endocytosis. *Neuron* **65**, 66–79 (2010).
  98. Arranz, A. M. *et al.* LRRK2 functions in synaptic vesicle endocytosis through a kinase-dependent mechanism. *J. Cell Sci.* **128**, 541–552 (2015).
  99. Munsie, L. N. *et al.* Retromer-dependent neurotransmitter receptor trafficking to synapses is altered by the Parkinson's disease VPS35 mutation p.D620N. *Hum. Mol. Genet.* **24**, 1691–1703 (2015).
  100. Deng, H. X. *et al.* Identification of TMEM230 mutations in familial Parkinson's disease. *Nature Genet.* **48**, 733–739 (2016).
  101. Satake, W. *et al.* Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nature Genet.* **41**, 1303–1307 (2009).
  102. Simón-Sánchez, J. *et al.* Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nature Genet.* **41**, 1308–1312 (2009).
  103. Ishizawa, T., Mattila, P., Davies, P., Wang, D. & Dickson, D. W. Colocalization of tau and  $\alpha$ -synuclein epitopes in Lewy bodies. *J. Neuropathol. Exp. Neurol.* **62**, 389–397 (2003).
  104. Cushman, M., Johnson, B. S., King, O. D., Gitler, A. D. & Shorter, J. Prion-like disorders: blurring the divide between transmissibility and infectivity. *J. Cell Sci.* **123**, 1191–1201 (2010).
  105. Jucker, M. & Walker, L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45–51 (2013).
  106. Clavaguera, F. *et al.* Brain homogenates from human tauopathies induce tau inclusions in mouse brain. *Proc. Natl Acad. Sci. USA* **110**, 9535–9540 (2013).
  107. Bae, E. J. *et al.* Glucocerebrosidase depletion enhances cell-to-cell transmission of  $\alpha$ -synuclein. *Nature Commun.* **5**, 4755 (2014).
  108. Walsh, D. M. & Selkoe, D. J. A critical appraisal of the pathogenic protein spread hypothesis of neurodegeneration. *Nature Rev. Neurosci.* **17**, 251–260 (2016).
  109. Kannarkat, G. T., Boss, J. M. & Tansey, M. G. The role of innate and adaptive immunity in Parkinson's disease. *J. Parkinsons Dis.* **3**, 493–514 (2013).
  110. Janeway, C. A. Jr & Medzhitov, R. Innate immune recognition. *Annu. Rev. Immunol.* **20**, 197–216 (2002).
  111. Czirr, E. & Wyss-Coray, T. The immunology of neurodegeneration. *J. Clin. Invest.* **122**, 1156–1163 (2012).
  112. Harms, A. S. *et al.* MHCII is required for  $\alpha$ -synuclein-induced activation of microglia, CD4 T cell proliferation, and dopaminergic neurodegeneration. *J. Neurosci.* **33**, 9592–9600 (2013).
  113. Su, X. *et al.* Synuclein activates microglia in a model of Parkinson's disease. *Neurobiol. Aging* **29**, 1690–1701 (2008).
  114. Kim, C. *et al.* Antagonizing neuronal Toll-like receptor 2 prevents synucleinopathy by activating autophagy. *Cell Rep.* **13**, 771–782 (2015).
  115. Lee, H. J., Suk, J. E., Bae, E. J. & Lee, S. J. Clearance and deposition of extracellular  $\alpha$ -synuclein aggregates in microglia. *Biochem. Biophys. Res. Commun.* **372**, 423–428 (2008).
  116. Zhang, W. *et al.* Aggregated  $\alpha$ -synuclein activates microglia: a process leading to disease progression in Parkinson's disease. *FASEB J.* **19**, 533–542 (2005).
  117. Conway, K. A., Rochet, J. C., Bieganski, R. M. & Lansbury, P. T. Jr. Kinetic stabilization of the  $\alpha$ -synuclein protofibril by a dopamine- $\alpha$ -synuclein adduct. *Science* **294**, 1346–1349 (2001).
  118. Martínez-Vicente, M. *et al.* Dopamine-modified  $\alpha$ -synuclein blocks chaperone-mediated autophagy. *J. Clin. Invest.* **118**, 777–788 (2008).
  119. Mosharov, E. V. *et al.* Interplay between cytosolic dopamine, calcium, and  $\alpha$ -synuclein causes selective death of substantia nigra neurons. *Neuron* **62**, 218–229 (2009).
  120. Maday, S., Wallace, K. E. & Holzbaur, E. L. Autophagosomes initiate distally and mature during transport toward the cell soma in primary neurons. *J. Cell Biol.* **196**, 407–417 (2012).
  121. Ashrafi, G., Schlehe, J. S., LaVoie, M. J. & Schwarz, T. L. Mitophagy of damaged mitochondria occurs locally in distal neuronal axons and requires PINK1 and Parkin. *J. Cell Biol.* **206**, 655–670 (2014).
  122. Chan, C. S. *et al.* 'Rejuvenation' protects neurons in mouse models of Parkinson's disease. *Nature* **447**, 1081–1086 (2007).
  123. Nixon, R. A. The role of autophagy in neurodegenerative disease. *Nature Med.* **19**, 983–997 (2013).
  124. Small, S. A. & Petsko, G. A. Retromer in Alzheimer disease, Parkinson disease and other neurological disorders. *Nature Rev. Neurosci.* **16**, 126–132 (2015).
  125. Schneider, J. L. & Cuervo, A. M. Autophagy and human disease: emerging themes. *Curr. Opin. Genet. Dev.* **26**, 16–23 (2014).
  126. Wong, Y. C. & Holzbaur, E. L. Autophagosome dynamics in neurodegeneration at a glance. *J. Cell Sci.* **128**, 1259–1267 (2015).
  127. Decressac, M. *et al.* TFEB-mediated autophagy rescues midbrain dopamine neurons from  $\alpha$ -synuclein toxicity. *Proc. Natl Acad. Sci. USA* **110**, E1817–E1826 (2013).
  128. Richter, F. *et al.* A GCase chaperone improves motor function in a mouse model of synucleinopathy. *Neurotherapeutics* **11**, 840–856 (2014).
  129. Daher, J. P. *et al.* Leucine-rich repeat kinase 2 (LRRK2) pharmacological inhibition abates  $\alpha$ -synuclein gene-induced neurodegeneration. *J. Biol. Chem.* **290**, 19433–19444 (2015).
  130. Mecozzi, V. J. *et al.* Pharmacological chaperones stabilize retromer to limit APP processing. *Nature Chem. Biol.* **10**, 443–449 (2014).
- The identification of a small molecule that increases the stability of the retromer complex and enhances its function, suggesting that such small molecules might also be effective at boosting retromer function in models of PD.**
131. Tran, H. T. *et al.*  $\alpha$ -Synuclein immunotherapy blocks uptake and templated propagation of misfolded  $\alpha$ -synuclein and neurodegeneration. *Cell Rep.* **7**, 2054–2065 (2014).
  132. Lazarou, M. *et al.* The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. *Nature* **524**, 309–314 (2015).
- Two receptors, optineurin and CALCOCO2 (also known as NDP52), were identified as important targets of PINK1 for recruitment to damaged mitochondria and for the initiation of selective autophagy (mitophagy).**
133. Südhof, T. C. The molecular machinery of neurotransmitter release (Nobel lecture). *Angew. Chem. Int. Edn Engl.* **53**, 12696–12717 (2014).
  134. Sanders, D. W., Kaufman, S. K., Holmes, B. B. & Diamond, M. I. Prions and protein assemblies that convey biological information in health and disease. *Neuron* **89**, 433–448 (2016).
  135. Stenmark, H. Rab GTPases as coordinators of vesicle traffic. *Nature Rev. Mol. Cell Biol.* **10**, 513–525 (2009).
  136. Luzio, J. P., Hackmann, Y., Dieckmann, N. M. & Griffiths, G. M. The biogenesis of lysosomes and lysosome-related organelles. *Cold Spring Harb. Perspect. Biol.* **6**, a016840 (2014).
  137. Lynch-Day, M. A., Mao, K., Wang, K., Zhao, M. & Klionsky, D. J. The role of autophagy in Parkinson's disease. *Cold Spring Harb. Perspect. Med.* **2**, a009357 (2012).
  138. Wong, Y. C. & Holzbaur, E. L. Optineurin is an autophagy receptor for damaged mitochondria in parkin-mediated mitophagy that is disrupted by an ALS-linked mutation. *Proc. Natl Acad. Sci. USA* **111**, E4439–E4448 (2014).

**Acknowledgements** We would like to thank S. Pfeffer for reading the manuscript. This work was funded by grants from the Michael J. Fox Foundation and the US National Institute of Neurological Disorders and Stroke.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: see [go.nature.com/2endued](http://go.nature.com/2endued). Readers are welcome to comment on the online version of this paper at [go.nature.com/2endued](http://go.nature.com/2endued). Correspondence should be addressed to A.A. (aa900@columbia.edu) or A.D.G. (agitler@stanford.edu).

**Reviewer Information** Nature thanks R. Nixon and the other anonymous reviewer(s) for their contribution to the peer review of this work.



# Mammalian prions and their wider relevance in neurodegenerative diseases

John Collinge<sup>1,2</sup>

**Prions are notorious protein-only infectious agents that cause invariably fatal brain diseases following silent incubation periods that can span a lifetime. These diseases can arise spontaneously, through infection or be inherited. Remarkably, prions are composed of self-propagating assemblies of a misfolded cellular protein that encode information, generate neurotoxicity and evolve and adapt *in vivo*. Although parallels have been drawn with Alzheimer's disease and other neurodegenerative conditions involving the deposition of assemblies of misfolded proteins in the brain, insights are now being provided into the usefulness and limitations of prion analogies and their aetiological and therapeutic relevance.**

Prions are lethal pathogenic agents that cause neurodegenerative diseases such as Creutzfeldt–Jakob disease (CJD) in humans, scrapie in sheep and goats, transmissible mink encephalopathy, chronic wasting disease in deer and elk, and bovine spongiform encephalopathy (BSE, or mad cow disease)<sup>1</sup>. According to the widely accepted ‘protein-only’ hypothesis, prions are devoid of nucleic acids and consist of multichain assemblies of misfolded, host-encoded prion protein (PrP), a cell-surface glycoprotein that is expressed in most tissues. Cellular PrP (PrP<sup>C</sup>) is present in all vertebrates and is highly conserved in mammals. Prions are thought to propagate by acting as a template or a seed that recruits PrP monomers to the multichain PrP assemblies, which are then elongated and fissioned in a physico-chemical process that mimicks the propagation of a biological pathogen. Not only do prions constitute infectious agents that are capable of entering and colonizing a host and evading its defences while being composed essentially of a single polypeptide, but the existence of several strains of prions that cause different patterns of disease constitutes protein-based inheritance; both aspects raise fascinating evolutionary questions. Similar protein-based inheritance is now well characterized in yeast and other fungi, and the field of yeast prions, which involves several different proteins, is advanced<sup>2</sup>.

Prion diseases in humans have three distinct causes: autosomal dominant inheritance (as a result of coding mutations in the PrP gene *PRNP*); spontaneous occurrence (that is, sporadic CJD), which is more common than inherited forms; or acquisition through environmental exposure to prions (via medical procedures, BSE-contaminated food or, formerly, endocannibalism in the case of the prion disease kuru in Papua New Guinea)<sup>1</sup>. Although this aetiological triumvirate was thought to be unique, its relevance to more common diseases is now being considered widely. Notably, the most common neurodegenerative diseases are associated with the accumulation of misfolded proteins and, similar to prion diseases, they also occur mainly as sporadic conditions but have rare, inherited forms associated generally with mutations in the genes that encode or process relevant accumulating proteins or peptides. Although these conditions were thought not to have acquired forms, this assumption is now being challenged through developing interest in whether prion-like mechanisms are involved, notably in Alzheimer's disease (AD) and Parkinson's disease (PD)<sup>3</sup>.

In this Review, my aim is not simply to consider the extent to which the basic concept of prion propagation — that is, seeded protein

aggregation — is relevant to these neurodegenerative conditions. I will also consider the wider understanding of prion pathogenesis before addressing the utility of this analogy. In this regard, it is essential to see prions as a dynamic collection or ‘cloud’ of misfolded protein assemblies that are maintained under biological selection in a host and its constituent tissues. Prions exist as distinct strains with an ability to adapt and mutate<sup>4–6</sup>, and with a kinetic and mechanistic relationship between propagating and neurotoxic species<sup>7</sup>. Such processes underlie the long, clinically silent incubation periods in prion diseases and the persistent carrier states that can span a lifetime and still pose risks to other people who are exposed to infected tissues<sup>8</sup>. These concepts from prion research might challenge conventional thinking and translational research strategies in neurodegeneration.

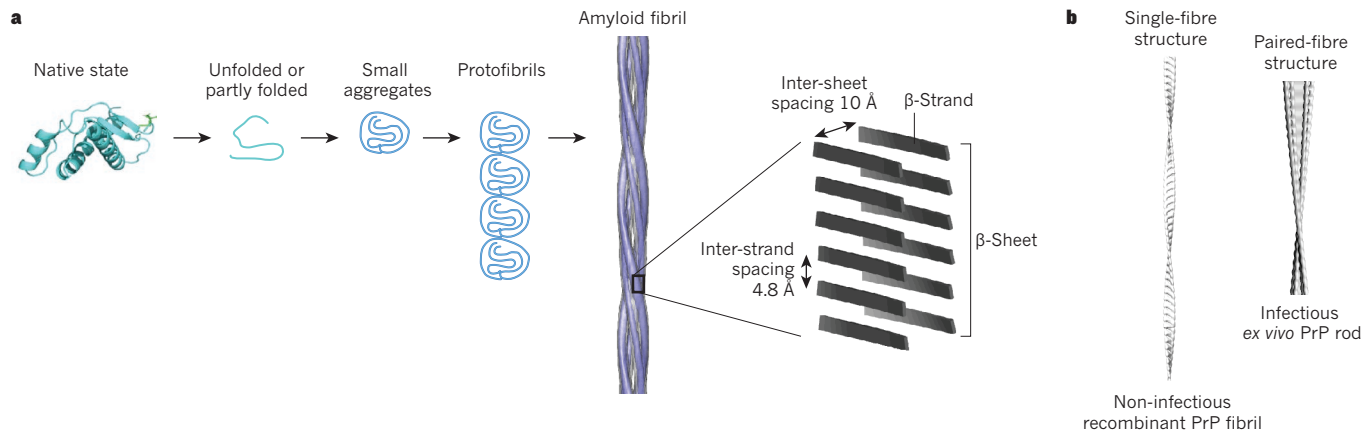
## The properties of prions

Although prion-like mechanisms are proposed in other neurodegenerative diseases, uncertainties still surround the structure and biological properties of mammalian prions. An improved understanding of prions and their propagation and pathogenicity should provide insights into the utility and wider relevance of prion analogies.

## Proteinaceous infectious agents

In the protein-only hypothesis, proposed in the 1960s (refs 9 and 10) and accepted widely since the 1990s, prions lack nucleic acids, are composed of misfolded isomers of PrP<sup>C</sup>, a normal constituent of the host brain, and propagate through the recruitment of PrP monomers<sup>11</sup>. The disease-associated isoform of PrP was designated as PrP<sup>Sc</sup>, and it was initially proposed that PrP<sup>Sc</sup> existed in a monomeric form and promoted the conformational change of PrP<sup>C</sup> with the aid of a hypothetical molecular chaperone known as ‘protein X’<sup>12</sup>. PrP<sup>Sc</sup> was defined biochemically in terms of its insolubility in detergents and its relative resistance to digestion by the protease proteinase K (PK)<sup>13,14</sup>; the term PrP<sup>Sc</sup> became used (and still is by many researchers) as synonymous with infectivity. Disease-associated PrP accumulates in various types of protein deposit in the brain, including amyloid plaques, and an alternative hypothesis — now accepted widely — proposes that prions consist of multimolecular assemblies of misfolded PrP monomers that form amyloids (Fig. 1a). Numerous proteins can adopt an amyloid state in which they form fibres with intertwined ribbons that consist of many stranded  $\beta$ -sheets (Fig. 1b), which are defined in biophysical terms by

<sup>1</sup>Medical Research Council Prion Unit, University College London Institute of Neurology, London WC1N 3BG, UK. <sup>2</sup>Department of Neurodegenerative Disease, University College London Institute of Neurology, London WC1N 3BG, UK.



**Figure 1 | Amyloids and prions.** **a**, Schematic steps in protein misfolding into amyloid. Many proteins seem able to form amyloid fibrils. These elongated fibres consist of long ribbons of  $\beta$ -sheet with the strands stacked perpendicular to the fibre axis. The pathway to formation of amyloid fibres is not understood fully but appears to involve nucleation from small aggregates that then assemble into protofibrils before developing into mature fibrils. Amyloid is defined pathologically by the demonstration of green birefringence on binding of the dye Congo red and biophysically by a cross- $\beta$  fibre diffraction

a cross- $\beta$  diffraction pattern. According to this hypothesis, prions are fibrillar forms of PrP, the ends of which constitute the infectious entity, and the exponential rise in the prion titre during an infection is a consequence of sequential elongation and fragmentation of the fibre<sup>15,16</sup>.

Although the multichain forms of PrP thought to comprise infectious prions are referred to generally as PrP<sup>Sc</sup>, it has been established that there are several disease-related forms of PrP, some of which are PK sensitive, and that these forms are responsible for most of the infectivity of some prion isolates<sup>17,18</sup>. At present, these species are poorly characterized in physical terms and a nomenclature has yet to be agreed. In this Review, the descriptive term 'PK-sensitive disease-related PrP' is used to refer to the ensemble of PrP species found in prion infections that do not meet the biochemical definition of PrP<sup>Sc</sup> (ref. 19).

### Prion strains and transmission barriers

Despite lacking a genome, mammalian prions exist as numerous strains that can be serially propagated in laboratory animals and that produce distinct patterns of disease<sup>20</sup>. Since several strains of prion can be maintained in mice from an inbred line with identical PrP genes, strains cannot be determined by differences in the amino-acid sequence of PrP. Strains are associated with biochemically distinct types of PrP<sup>Sc</sup> that are thought to represent structurally distinct seeds that are able to recruit the host PrP<sup>C</sup> into their distinct multimolecular assemblies<sup>21–24</sup>. These strain-specific biological and structural properties can be maintained even after passage through an intermediate mammalian species with a different PrP amino-acid sequence<sup>22,25</sup>.

Although prions from one species of mammal can infect another, in practice this is much less efficient than transmission between individuals from a single species. This is known as the species barrier or the transmission barrier and was initially thought to come from the degree of amino-acid sequence homology between the PrPs of the two species involved<sup>26</sup>. However, it became clear that the prion strain was crucial as two strains propagated in identical hosts may have quite different transmission barriers to a second species. A dramatic example of this is provided by a study of variant CJD (vCJD), a disease caused by exposure to BSE prions. Prions from people with sporadic CJD and an identical *PRNP* genotype can be readily transmitted to transgenic mice that express only human PrP but not at all to wild-type mice<sup>27</sup>. However, the transmission of prions from people with vCJD to wild-type mice occurred with greater efficiency than their transmission to transgenic mice with PrP characteristics similar to those of prions

pattern. **b**, Models of recombinant PrP fibrils and *ex vivo* prions. Amyloid, generally manifest as extracellular plaques, is often but not always apparent in prion diseases. Fibres that meet the tinctorial and biophysical definition of amyloids can be readily produced from recombinant prion proteins that are not infectious in cellular or animal models<sup>69</sup> (left). Authentically infectious, *ex vivo* prions have a more complex assembly state consisting of rods formed from short, paired double-helical fibrils that have so far not been produced synthetically<sup>85</sup> (right).

from BSE-affected cattle<sup>22</sup>. Such findings led to the development of the conformational selection model<sup>28,29</sup> (Fig. 2) in which only a subset of all possible PrP assemblies that can effectively propagate in mammalian tissues are compatible with a particular host's PrP<sup>C</sup> sequence and can act as pathogens in the host. Transmission barriers can be explained by the degree of overlap between the permissible conformers between the two species involved (Fig. 2a). Under this model, the barriers that control the transmission of prions between species are fundamentally related to strains and could be considered as opposite sides of the same coin. Such transmission barriers and the effects of conformational selection might also be encountered within a single species of mammal in which there are PrP coding polymorphisms, such as that at amino-acid residue 129 in humans that results in the presence of either methionine or valine and has a powerful effect on susceptibility and strain selection<sup>30</sup>.

### Quasispecies and prion strain mutation

The phenomenon of prion strain mutation, in which a strain does not breed true on passage in a new host and a distinct strain is propagated instead, has long been recognized using biological methods (ref. 4). Such mutation might occur when prions infect a new species but it can also occur in the same species when the PrP amino-acid sequence of the inoculated prions differs from that of the host<sup>22,30,31</sup> (Fig. 2b). Strain mutation can be readily accommodated by the conformational selection model<sup>5,29</sup>. Although prion strains can be biologically cloned<sup>32</sup>, some strains are intrinsically unstable<sup>21</sup> and revert readily to another strain, and natural isolates can contain several strains or types of PrP<sup>Sc</sup> (refs 22–35). Prion isolates show considerable heterogeneity when assessed using a range of biochemical and physical methods; there is marked diversity in N-terminal protease cleavage sites, which indicates conformation diversity, and PrP glycosylation is complex and highly variable. They also show a heterogeneity in terms of heat inactivation, with the presence of thermostable subpopulations<sup>36</sup>. The proposal that prion strains, instead of being clonal, actually constitute a cloud of diverse molecular assemblies, which is analogous to a viral quasispecies<sup>29</sup>, is now supported strongly by evidence showing that biologically cloned prion strains can adapt under selection pressure in cell culture<sup>6,37</sup>.

Prion strains may therefore be considered as subtypes of misfolded PrP assemblies with the necessary properties and replication kinetics to evade cellular and other host defences, propagate exponentially and be able to act as an efficient pathogen while other PrP assemblies are



degraded or form highly stable aggregates that do not propagate or are comparatively inert<sup>29</sup>. The very diversity of prion quasiespecies, the degree of which may itself vary between strains, could be fundamental to the adaptation of a strain to a new host (and to particular tissues in it) and to enable its survival and spread. This might also account for why some amyloids are infectious pathogens and others are not.

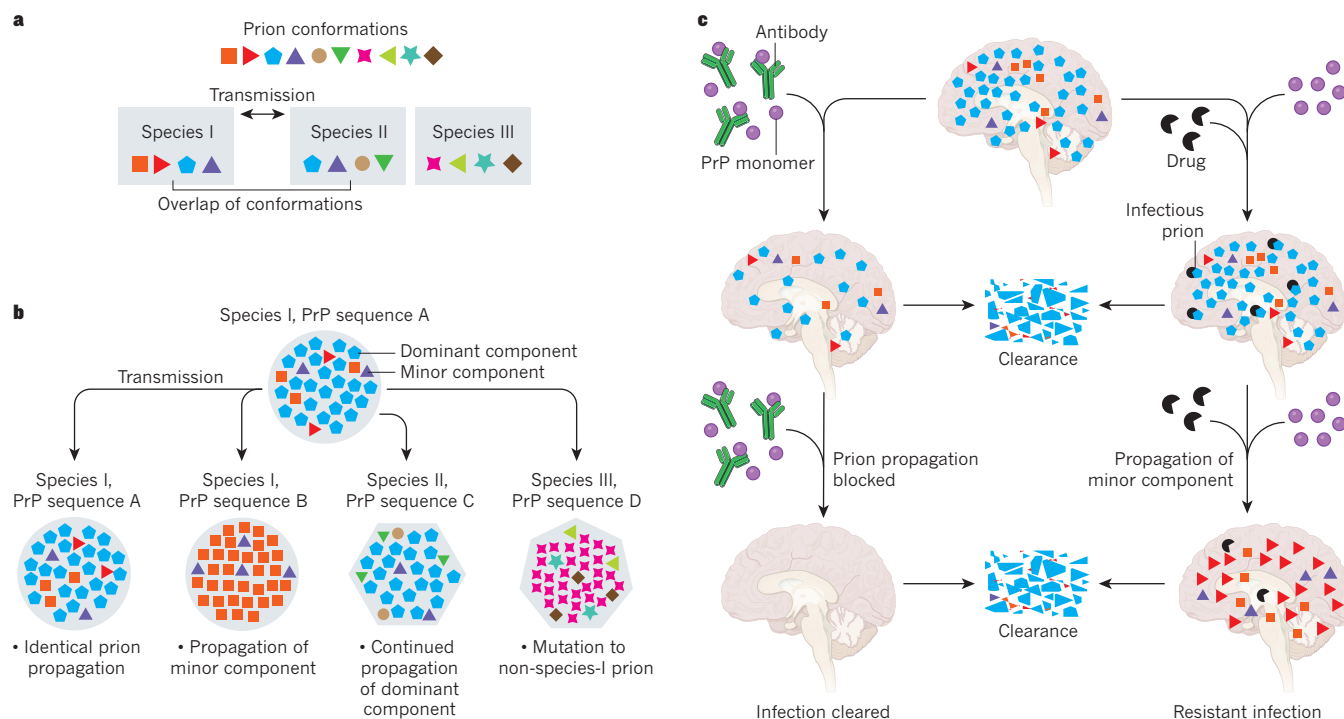
### The kinetics of prion propagation and neurotoxicity

Despite the proposal of various mechanisms<sup>38–42</sup>, the cause of cell death in prion-related neurodegeneration remains unclear. Although the PrP gene is highly conserved, the loss of PrP<sup>C</sup> function is not sufficient to cause cell death<sup>43,44</sup>, although the expression of PrP is an obligate requirement for prion propagation and pathology<sup>45–48</sup>.

Prion infections are invariably associated with clinically silent incubation periods that can exceed 50 years in people<sup>49</sup> and yet are followed by an aggressive clinical phase that is typically measured in months<sup>1</sup>. In inbred mouse lines, incubation periods for a defined strain of mouse prion are remarkably reproducible, with the onset of clinical symptoms in a group of intracerebrally inoculated mice spanning only a few days following incubation periods of six months or more. The length of these incubation periods is known to correlate inversely with the level of PrP<sup>C</sup> expression<sup>26</sup>. The development of cell-based prion bioassays have enabled more detailed and precise measurement of the levels of infectivity throughout the incubation period and their correlation with PrP expression. Surprisingly, this demonstrated that prion propagation in the brain proceeds in two distinct mechanistic phases (Fig. 3). The first

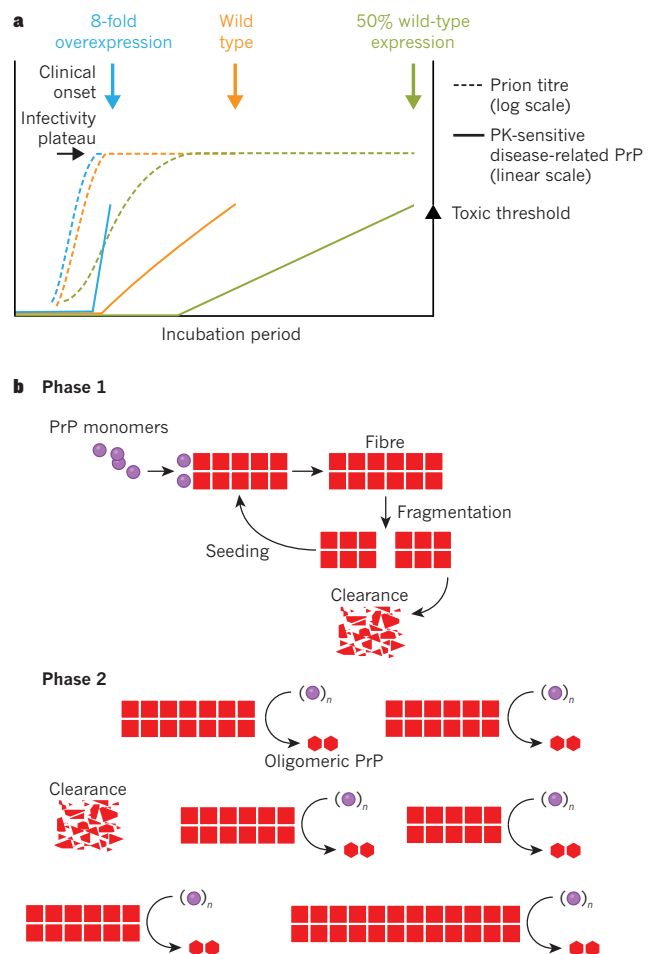
is a clinically silent exponential phase, the rate of which is not limited by PrP<sup>C</sup> concentration, that rapidly reaches a maximal prion titre. This is followed by a phase that is characterized by plateauing of the prion titre, which continues until the onset of clinical signs. The length of the second phase is inversely proportional to the level of PrP expression, which indicates that the nature of the conversion process has changed (Fig. 3) as it must be rate limited by the addition of PrP<sup>C</sup>. Prion propagation and neurotoxicity can therefore be uncoupled, which supports the argument that prions are not directly neurotoxic. It is hypothesized that once prion propagation saturates, there is a switch in mechanism from the autocatalytic production of infectivity (phase 1) to a toxic pathway in which prions act as a catalytic surface for the production of toxic species (designated PrP<sup>L</sup>, in which 'L' stands for lethal) through a process that is now linearly dependent on PrP<sup>C</sup> concentration<sup>7</sup> (phase 2).

The measurement of both classic PrP<sup>Sc</sup> and PK-sensitive disease-related PrP isoforms during their incubation in mice with different levels of PrP expression demonstrates a linear increase in PK-sensitive disease-related PrP isoforms, distinct from classic PrP<sup>Sc</sup>. The increase commences at the phase transition and its rate is proportional to the PrP<sup>C</sup> concentration (Fig. 3a). The clinical onset of symptoms occurs at a similar level of PK-sensitive disease-related PrP isoforms, irrespective of the PrP<sup>C</sup> expression level, and the classic neuropathological features of prion disease, which indicate neurotoxicity, are detected only after the second phase has been well established. It is unknown whether most of the PK-sensitive disease-related PrP isoforms are toxic or — at the other extreme — whether a low-abundance PrP subspecies of defined



**Figure 2 | Prion propagation.** **a**, The transmission of prions is controlled by a mechanisms known as conformational selection<sup>28,29</sup>. A limited number of misfolded mammalian PrP conformations can propagate as prions *in vivo* and only a subset of these are compatible with each mammalian PrP<sup>C</sup> sequence. Prions transmit readily between individuals of the same species that express the same PrP<sup>C</sup> and might transmit between species with different PrP<sup>C</sup> sequences (for example, species I and species II) if there is an overlap of permissible conformations. When there is little or no overlap, such as between species I and species III, a considerable transmission barrier occurs and transmission might require strain mutation. Transmission barriers between individuals of the same species can be introduced by PrP polymorphisms. **b**, Prion strains are not clonal but constitute a molecular ensemble that is maintained under host selection with a dominant component and an array of minor components<sup>5,6</sup>. A strain can propagate faithfully in an individual of the

same species that expresses an identical PrP<sup>C</sup> (species I, PrP sequence A), or a minor component of the strain might be selected from the ensemble when prions colonize an individual from the same species that expresses a different PrP<sup>C</sup> (species I, PrP sequence B). A strain might also propagate faithfully in a different species that has a compatible PrP<sup>C</sup> (species II, PrP sequence C). If the strain is not compatible with a separate species (species III, PrP sequence D), its mutation to a type that cannot be detected in the original ensemble must have occurred to establish an infection. **c**, Prion propagation requires a supply of PrP monomers (purple spheres). Using a drug that binds to infectious prions increases the risk of the rapid development of resistance because structurally diverse conformers, which are present as minor components, are selected for propagation<sup>37</sup>. A drug or antibody that binds to native PrP<sup>C</sup> monomers should, in principle, block the propagation of all prion conformers, which can then be cleared by existing host mechanisms<sup>7,90</sup>.



**Figure 3 | Kinetics of prion propagation and toxicity.** **a**, The propagation of Rocky Mountain Laboratory (RML) prions during the incubation period in congenic mice is shown for three levels of PrP<sup>C</sup> expression: wild-type expression (orange), eight-fold overexpression (blue) and 50% of wild-type expression (green). Prion replication is exponential in the clinically silent phase 1 until a limiting titre of prions (infectivity plateau) is rapidly attained. This is the same for all three mouse lines and is not rate limited by the level of PrP<sup>C</sup> expression. Phase 2 follows, in which the plateau prion titre continues until the clinical onset of neurological disease (vertical arrows). Its length is inversely proportional to the PrP<sup>C</sup> expression level<sup>7</sup>; the nature of the conversion process has changed because it is now rate limited by the addition of PrP<sup>C</sup>. At the phase transition, when the infectivity plateau is reached, the levels of PK-sensitive disease-related PrP isoforms begin to rise in a linear fashion that is now proportional to the corresponding level of PrP<sup>C</sup> expression. Classic (PK-resistant) PrP<sup>Sc</sup> levels (not shown) represent only a small fraction of total disease-related PrP and the majority of this is detected at the end of phase 2 (ref. 19). Clinical onset occurs at a similar level of PK-sensitive disease-related PrP isoforms, irrespective of PrP<sup>C</sup> expression level, which suggests that there is a common toxic threshold (black triangle). Classic neuropathological features of prion disease can be detected only after plateauing has been well established<sup>19</sup>. Unaltered levels of disease-related PrP isoforms during the exponential phase, when prion titres rise by a million fold, indicate that prions form a small minority of the total PrP. **b**, After the propagation of prions becomes saturated, there is a mechanistic switch from the autocatalytic production of infectivity (phase 1, top) to a toxic pathway (phase 2, bottom) in which prion particles act as a catalytic surface for the production of toxic species presumed to be oligomeric forms of PrP (PrP<sup>L</sup>) in a process that is now linearly dependent on the PrP<sup>C</sup> concentration<sup>7,29</sup>.

structure and receptor interaction with high specific toxicity constitutes PrP<sup>L</sup>. These findings also suggest that the term PrP<sup>Sc</sup>, which is often used as a synonym for infectivity, should be restricted to describe material that meets its original biochemical definition. Such classic PrP<sup>Sc</sup> forms

only a small and variable minority of the total disease-related PrP isoforms and its relative contribution to both infectivity and toxicity is unclear.

Alternative models can also be envisaged. For example, the fission of infectious particles could cease, leading to their further growth but without an increase in the number of particles overall, or a crucial cellular component might be depleted<sup>7</sup>. It has also been proposed that PrP<sup>C</sup> acts a receptor for PrP<sup>Sc</sup> and mediates toxic signalling<sup>38</sup>. However, this is challenged by the existence of subclinical prion carrier states in which mice with normal neuronal PrP<sup>C</sup> expression do not develop clinical symptoms and experience a normal lifespan despite having PrP<sup>Sc</sup> levels (and prion titres) that are as high as those of end-stage, clinically ill mice<sup>8,50–53</sup>.

### Prion spread

The term prion-like spread is increasingly being used with respect to the evolution of pathology in the more common neurodegenerative conditions, notably AD and PD, to imply the cell-to-cell transmission of disease-related amyloid seeds and their spread along neuroanatomical pathways<sup>3,54,55</sup>. However, although prions have been shown to traverse these pathways, including the optic pathway<sup>56</sup>, they characteristically target highly defined regions or cell layers of the brain, which can occur irrespective of the anatomical route or physical location on the animal of the prion inoculation site<sup>57</sup>. Indeed, such specific and remarkably reproducible neuroanatomical targeting is how distinct strains of prion were identified originally and are still distinguished<sup>58</sup>. Although proteopathic seeds, including prions, would be expected to follow neuroanatomic pathways to some degree with diffusion along tissue planes, and because of the cell-to-cell contact and the vesicular transport that is inherent at synapses, for example, the use of the description ‘prion-like’ could be misleading. This is because the term implies specific neuroanatomical targeting that is independent of the site at which the seeding process is initiated. Indeed, this term might be even more questionable because the classic methods used to distinguish prion strains are based on neuropathology, and not prion titres, in specific regions of the brain. The actual spread of prions could therefore be different to the spread of the pathology that is now known to develop only after prion levels have saturated in whole brain<sup>19</sup>, suggesting that this may relate to selective cellular vulnerability. Automated cell-based prion bioassays<sup>59</sup> may allow fine mapping of the spread of prions as the brain is colonized, which would facilitate comparative studies with mouse models of seeding of other neurodegenerative pathologies to investigate the extent to which these were actually prion-like. Prototype *in vitro* amplification assays have been reported for amyloid- $\beta$  seeding activity and for the protein  $\alpha$ -synuclein<sup>60,61</sup>, although these would need to be validated against *in vivo* or cell-based propagation assays.

### The structure of infectious prion particles

It has proved challenging to adequately purify *ex vivo* prions, and the material obtained is both insoluble and heterogeneous, which renders it unsuitable for high-resolution structural methods such as X-ray crystallography or multidimensional nuclear magnetic resonance spectroscopy. Importantly, given this heterogeneity and the effects of processing on infectivity, it has been difficult to unequivocally correlate infectivity with structural features imaged in previous studies<sup>62,63</sup>. The structure of infectious particles and the structural basis of prion-strain diversity therefore remain unresolved.

An alternative approach has focused on generating mammalian prions from bacterially expressed recombinant PrP (rPrP). Success would place the protein-only hypothesis beyond doubt and also provide a model system in which the mechanism of assembly and the structural basis of prion infectivity could be established. Early studies reported the conversion of the predominantly  $\alpha$ -helical PrP<sup>C</sup> conformation to isoforms with the biochemical features of *ex vivo* PrP<sup>Sc</sup>; however, these isoforms were not associated with detectable infectivity<sup>64–68</sup>. It is clear that the production of amyloid fibrillar forms of rPrP is not sufficient to generate mammalian prions with infectivity that can be demonstrated



through bioassays<sup>69</sup>. Furthermore, prion disease has been transmitted in the absence of detectable PrP amyloids or PrP<sup>Sc</sup> (refs 70 and 71). Subsequent studies reported the *de novo* production of prions from rPrP, each using different methods<sup>72–76</sup>. The interpretation of such findings is complicated by the use of bioassays in transgenic mice that develop spontaneous neurodegeneration or *in vitro* methods of prion amplification such as protein misfolding cyclic amplification, which can amplify a single prion particle to a level detectable by bioassay; these risk the amplification of low-level contamination or prions that form spontaneously in the brain at low frequency<sup>77</sup>.

Given Avogadro's number, the production of relatively homogenous infectivity from purified rPrP would be expected to result in extremely high specific infectivity that could be diluted many millions of times and still produce 100% lethality in a rodent bioassay. Without such very high titres, structural studies would only report on the large excess of non-infectious material<sup>29</sup>. Crucially, none of the methods for production of synthetic prions published so far enable the systematic production of high-titre material that is suitable for structural studies. A high-throughput cell-based prion bioassay has been used to systematically screen around 20,000 unique experimental conditions to generate prions from rPrP in a manner that is conceptually similar to a matrix approach to protein crystallization<sup>59</sup>. Although some seemingly infected cell cultures were obtained, these were transient and could not be reproduced reliably. It is possible therefore that efficient prion production requires post-translationally modified PrP, specific co-factors or biological replication sites or surfaces<sup>59</sup>. One possibility is that only a defined subset of PrP assemblies with highly specific quaternary structural features might be capable of acting as effective *in vivo* pathogens (with a propagation rate that exceeds the capacity of the host's clearance mechanisms). These PrP assemblies would form at an extremely low probability *in vitro* but could then be selected and amplified in a biological system. The systematic production of high-titre prions from rPrP therefore probably requires a detailed knowledge of the quaternary structure of authentic *ex vivo* prions.

Historically, the most highly enriched preparations of mammalian prions have contained a large excess of PrP per infectious particle<sup>78–80</sup>, leading to prolonged debate about whether only a minority of PrP molecules are infectious or whether PrP aggregates of a defined size are necessary for infectivity<sup>80–82</sup>. Although the idea that mammalian prions are self-propagating fibrillar or amyloid forms of PrP (refs 15, 16 and 29) is appealing and can readily accommodate observed variance in specific infectivity with respect to PrP monomers, historically PrP assemblies known as prion rods<sup>78</sup> have been described as being formed *in vitro* during PrP purification by a process that requires both detergents and limited proteolysis<sup>83</sup>. There has been long-standing uncertainty as to whether purified PrP structures that are associated with infectivity have authentic, biologically relevant morphologies.

Methodological advances that enable the production of *ex vivo* prions of exceptionally high purity and *in situ* assay of infectivity on electron microscope grids have shown that rod-like assemblies can be isolated from brain tissue in the absence of detergent and without treatment with PK; consequently, the physical relationship between such structures and infectivity has been established<sup>84,85</sup>. These infectious PrP rods, isolated from several prion strains, have a common hierarchical assembly that involves twisted pairs of short fibres with a repeating substructure (Fig. 1b). The properties that define prion strains are retained by such PrP polymer preparations, which indicates that they have an authentic prion composition and structure<sup>85</sup>. This specific 3D structure, which is distinct from the long, single fibres that make up non-infectious fibrils prepared from rPrP (ref. 69), provides a basis for understanding the features that discriminate prions from non-infectious protein assemblies. The paired-fibre structure might contribute considerably to the resistance of prions to degradation, which is probably crucial to their ability to act as transmissible agents<sup>85</sup>. The characteristic gap of 8–10 nanometres between the two fibres in the rods might be occupied by *N*-linked glycans that could contribute to the stability of the assembly, and consequently

its infectivity, as well as provide a steric basis for the consistent ratios of PrP glycoforms that characterize distinct prion strains<sup>22,86</sup>.

### Prion therapeutics

Prion diseases are invariably lethal conditions with generally rapid progression and no disease-modifying treatments. PrP<sup>C</sup> is an attractive therapeutic target because it is the obligate substrate for the production of all propagating or toxic disease-related PrP assemblies. The constitutive or conditional knockout of PrP<sup>C</sup> expression has no major phenotype<sup>43,44,87</sup> and targeting its expression in adult neurons during neuro-invasive prion disease prevents clinical onset and leads to the reversal of early spongiform neuropathology and behavioural deficits in mice<sup>48,88</sup>. Because PrP<sup>C</sup> needs to unfold to adopt the  $\beta$ -sheet-rich structure of the infectious amyloid form<sup>89</sup>, the binding of a ligand to the folded domain of PrP reduces the availability of unfolded PrP for prion propagation by acting as a pharmacological chaperone<sup>90</sup>. Conversely, agents that bind to disease-related PrP have resulted in the rapid development of drug resistance through strain mutation and the selection of propagating conformers<sup>6,37,91</sup> (Fig. 2c). Although it has proved challenging to develop drug-like small-molecule PrP<sup>C</sup> ligands, anti-PrP monoclonal antibodies cure prion-infected cells readily<sup>92–94</sup> and passive immunotherapy with monoclonal antibodies that bind PrP<sup>C</sup> has shown powerful therapeutic effects in several animal models, with clinical studies under development<sup>95–98</sup>.

### Misfolded proteins in other neurodegenerative diseases

The accumulation of multimolecular assemblies of various misfolded proteins in the brain is a long-established pathological feature of several neurodegenerative diseases. Much research has focused on the ability of such assemblies to propagate and spread in experimental models, with considerable potential implications for the aetiology, prevention and treatment of these diseases.

### The transmissibility of other neurodegenerative diseases

The landmark demonstrations of transmissibility through intracerebral inoculation with the homogenates of brains affected by the degenerative neurological disorder kuru and then CJD to chimpanzees in the 1960s led to the recognition of human transmissible spongiform encephalopathies, now known as prion diseases<sup>99,100</sup>. This was followed in the 1970s by attempts to determine whether other neurodegenerative diseases, including AD and PD, might also be transmissible to primates. Despite the inoculation of hundreds of animals and their observation for many years, no clear evidence of transmission was obtained<sup>55,101</sup>. However, in later studies in which marmosets were intracerebrally inoculated with AD brain homogenate, although no overt clinical disease was noted, several animals were found to have the amyloid- $\beta$  (but not tau protein) pathology that is characteristic of AD in autopsies carried out 6–7 years later<sup>102,103</sup>. AD is characterized pathologically by the accumulation of extracellular amyloid- $\beta$  plaques and vascular deposits of amyloid- $\beta$ , a condition termed cerebral amyloid angiopathy (CAA), as well as intracellular neurofibrillary tangles composed of assemblies of the microtubule-associated protein tau. Amyloid- $\beta$  peptides are a normal cleavage product of the amyloid precursor protein (APP) and there is persuasive genetic and biomarker-based evidence to support amyloid- $\beta$  as the root cause of AD<sup>104</sup>. Amyloid- $\beta$  is prone to self-association and aggregation and exists in numerous assembly forms. Soluble, oligomeric forms of amyloid- $\beta$  have a higher neurotoxicity than either monomeric amyloid- $\beta$  or mature amyloid- $\beta$  fibrils<sup>105,106</sup>. These soluble oligomers vary in size from dimers to complex structures that contain hundreds of peptides in different morphologies<sup>107</sup>.

There are now numerous reports that the production and spread of amyloid- $\beta$  deposits can be triggered in appropriate transgenic mouse models by inoculation with brain extracts from AD patients. Although rodent amyloid- $\beta$  does not readily form aggregates *in vivo*, transgenic mice that express mutant human APP spontaneously develop both amyloid- $\beta$  plaques and CAA as they age but this occurs much earlier

following intracerebral injection with extracts of amyloid- $\beta$ -containing AD brains<sup>108,109</sup>. This induction of amyloid- $\beta$  pathology is dependent on the concentration of amyloid- $\beta$  in the brain extract and can be blocked by active or passive immunization of the mice against amyloid- $\beta$ <sup>109</sup>. Subsequently, synthetic assemblies of amyloid- $\beta$  have also been shown to accelerate the development of amyloid- $\beta$  pathology in these mice<sup>110</sup>. Importantly, inoculation with AD brain extracts can induce amyloid- $\beta$  pathology in transgenic rodents that do not spontaneously develop such pathology<sup>111</sup>. Intraperitoneal inoculation with AD brain extracts also accelerates the onset of amyloid- $\beta$  pathology in transgenic mice and, interestingly, such peripherally seeded mice tend to develop mostly CAA rather than the parenchymal amyloid- $\beta$  deposits that are seen in these mice as they age<sup>112</sup>. These findings suggest that mechanisms exist for the transport of amyloid- $\beta$  seeds — as well as other proteopathic seeds such as tau<sup>113</sup> — from the periphery to the brain.

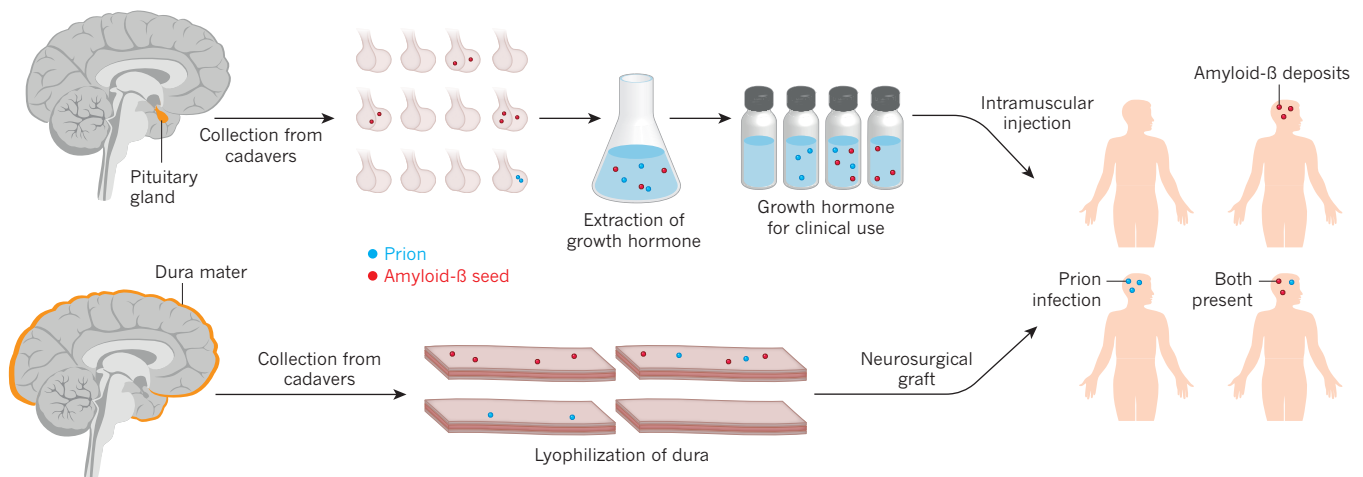
These animal studies provide convincing evidence that amyloid- $\beta$  pathology can be seeded by inoculation. However, a crucial caveat is that the animals do not develop progressive neurodegeneration, which calls into question the relevance of these findings for the naturally occurring human disease. By definition, amyloids should seed and propagate in the presence of a suitable concentration of monomer and a permissive environment — this is simply a thermodynamically driven physico-chemical process. Yet these conditions are not sufficient for the development of the full, lethal neurodegenerative syndrome. Inoculation with brain homogenates from people with the neurodegenerative condition multiple system atrophy, in which  $\alpha$ -synuclein forms intracellular amyloid deposits, was reported to produce a neurodegenerative illness in transgenic mice that are hemizygous for a transgene expressing a mutant  $\alpha$ -synuclein. However, homozygous transgenic mice expressing the same mutant protein develop spontaneous disease, and wild-type mice or those that express a non-mutant  $\alpha$ -synuclein show no disease when similarly inoculated with tissue affected by multiple system atrophy<sup>114</sup>.

### The relevance of seeding to naturally occurring disease

Of paramount importance is whether such experimental seeding of amyloid deposits in animal models is relevant to neurodegenerative diseases in humans. In particular, might AD be transmissible between

humans? Although there is no epidemiological evidence to support this possibility, it should be remembered that human prion diseases such as CJD, unequivocally a transmissible disorder, do not pass from person to person even with prolonged intimate contact. Instead, they require inoculation during medical or surgical procedures, which causes iatrogenic CJD (iCJD). Recognized routes to transmission include: injection with pituitary hormones derived from human cadavers, the grafting of human tissues (such as the dura mater or the cornea) and intracerebral exposure to contaminated neurosurgical instruments. There is also epidemiological evidence to suggest that common general surgical procedures might transmit CJD<sup>115,116</sup>. Evidence now suggests that such seeding of amyloid- $\beta$  pathology has occurred in humans. Before 1985, children with growth deficiencies were treated with intramuscular injections of growth hormone extracted from large pools of cadaver-derived pituitary glands (c-hGH), and a small proportion of these individuals have developed CJD after long incubation periods (Fig. 4). Around 30,000 people worldwide are thought to have received such treatments.

Autopsies of eight people with c-hGH-related iCJD and very long incubation periods found that four also had moderate to severe parenchymal and vascular amyloid- $\beta$  pathology typical of that seen in AD and CAA that did not colocalize with PrP deposition, and a further two people had focal amyloid- $\beta$  pathology<sup>117</sup>. The people studied were relatively young (aged 36–51) and had been treated with c-hGH during childhood, around 30 years earlier. Such amyloid- $\beta$  pathology is extremely rare in this age group and was not seen in people of up to a decade older who had died of prion diseases unrelated to c-hGH therapy. Amyloid- $\beta$  deposits are found in the pituitary glands of people with parenchymal amyloid- $\beta$  pathology<sup>117,118</sup>, and it therefore seems that some c-hGH preparations were probably contaminated with amyloid- $\beta$  seeds as well as CJD prions, leading to the simultaneous iatrogenic transmission of amyloid- $\beta$  pathology. Although none of the eight people studied had the tau neurofibrillary tangles that also characterize AD, these features are thought to appear later in the evolution of AD, and it is possible the full neuropathology of AD would have developed had these people not succumbed to CJD at a young age. However, the vascular amyloid- $\beta$  deposition seen unequivocally met the criteria for CAA, meaning that



**Figure 4 | Transmission of amyloid- $\beta$  pathology in humans.** Children with a growth deficiency caused by one of several conditions were treated with growth hormone extracted from large pools of pituitary glands obtained from cadavers (top). Although this practice ceased in 1985, when the risk of transmitting CJD was recognized, people are still developing iCJD because of the long incubation periods of prion infections in humans. A 2015 study of people who had recently developed iCJD found that most also had considerable deposits of amyloid- $\beta$  in their brain tissue and notably around the blood vessels of the brain (a condition known as CAA)<sup>117</sup>. Such features are characteristic of AD but are rarely present at the relatively young ages of the iCJD cohort, which suggests that the growth hormone they received might

have been contaminated by amyloid- $\beta$  seeds as well as by prions. However, the iCJD cohort did not have tau protein aggregates that are also characteristic of AD and it is unknown if these would have developed had the patients lived for longer. However, several individuals were developing CAA, which would have led to cerebral haemorrhages. The dura mater, a tough membrane that surrounds the brain, was also sourced from cadavers and processed before being used as grafts in neurosurgical procedures (bottom). This led to many occurrences of iCJD worldwide, and such use of dura mater has now been discontinued. Similar findings to those reported from treatment with growth hormone are consistent with the hypothesis that the amyloid- $\beta$  deposits were seeded from contaminated dura mater grafts<sup>119,120</sup>.

these individuals would probably have developed cerebral haemorrhages had they lived longer. This study raises the possibility that other known routes to prion transmission might be relevant to amyloid- $\beta$ , and similar findings have now been reported for iCJD related to dura mater grafting<sup>119,120</sup> (Fig. 4).

Although CJD is not contagious, this characteristic could be related to the tissue distribution of prions, which are confined mostly to the central nervous system (CNS)<sup>101</sup>. This is also true for kuru, meaning that its dietary transmission probably requires the ingestion of CNS tissue<sup>121</sup>. However, the tissue distribution in vCJD, which is caused by a different strain of prion, is much wider and shows extensive colonization of lymphoreticular tissues<sup>122</sup>; vCJD infection has been transmitted by blood products<sup>123</sup>. Some animal prion diseases, however, are clearly contagious. Scrapie spreads in flocks of sheep and is notoriously hard to eradicate and, most strikingly, chronic wasting disease spreads rapidly in free-ranging wild deer and elk, achieving a remarkably high prevalence in some areas of the United States. Chronic wasting disease prions are detected readily in the urine, faeces and saliva of affected deer<sup>124</sup>. The tissue distribution of prions is determined by a combination of host factors and the prion strain involved<sup>1</sup>. It will be important to determine the tissue distribution of amyloid- $\beta$  and other proteopathic seeds to establish whether risk of transmission is confined to the use of surgical instruments, biological medical products and grafts involving CNS tissue or whether there are wider potential risks, notably from blood transfusion or general surgery. This Review focuses on comparisons with AD but similar concerns relate to other neurodegenerative diseases, including those related to the accumulation of tau and  $\alpha$ -synuclein<sup>54</sup>. These are intracellular proteins, however, and their accumulation might involve mechanisms that are distinct from those that facilitate the seeding of prions, amyloid- $\beta$  assemblies and systemic amyloids.

### Strains of amyloid- $\beta$ and other proteopathic seeds

Evidence for the existence of strains in other neurodegenerative diseases has also emerged from studies that correlate molecular structural variations in protein assembly states with distinct biological activities or phenotypes of patients<sup>125–131</sup>. The experimental seeding of amyloid- $\beta$  pathology in transgenic mice suggests that several strains of amyloid- $\beta$  might contribute to phenotypic variability in AD<sup>109,132</sup>. Biological analysis of prion strains has relied on the ease of transmission of rodent-adapted scrapie prions to multiple inbred mouse lines<sup>25</sup>. The replacement of murine with human PrP<sup>C</sup> expression in transgenic mice has enabled similar studies of human prion strains<sup>133</sup>. These studies involve the transmission of prions to animals that do not spontaneously develop amyloid pathology or neurodegeneration; interpretation of strain-typing studies is known to be complicated by the expression of mutant PrP (ref. 134). Transmission studies of amyloid- $\beta$  and other proteopathic seeds have so far relied on transgenic models of the respective inherited forms of the neurodegenerative diseases that not only develop spontaneous pathology but might also show modified strain characteristics when non-wild-type substrate proteins are expressed, as seen in prion models<sup>134</sup>. Even more challenging for isolating and understanding the importance of strains of proteopathic seeds in AD is the fact that two proteins are polymerizing, amyloid- $\beta$  and tau, and the temporal and mechanistic relationship between these two processes remains unclear<sup>135</sup>. Furthermore, different lengths and isoforms exist of both amyloid- $\beta$  and tau proteins<sup>54</sup>. Amyloid- $\beta$  peptides are produced from APP through sequential endoproteolytic cleavages that can occur at several sites, producing peptides with a range of lengths and differing propensities for aggregation. As a result of alternative mRNA splicing, tau occurs in six isoforms — all of which are found in the tau filaments of AD. To a mammalian prionologist, aware of the complexity of characterizing and understanding strains composed of a single PrP species, unravelling such complexity seems a daunting task. Strains of amyloid- $\beta$  and tau might be important pathogenetic variables. If so, and by analogy with the wide phenotypic variation seen in prion diseases in humans, such strains — and their possible

interactions — might readily generate a diversity that encompasses the phenotypic heterogeneity of AD and other diseases that are associated with amyloid- $\beta$  and tau neuropathology.

### Implications of prion-like mechanisms for therapeutic agents

Although no treatments exist at present to halt the progression of AD, or indeed any other neurodegenerative diseases, there are many potential therapeutic approaches<sup>135</sup>. If prion-like processes do contribute to the pathogenesis of AD, what are the implications for research into therapeutic drugs for AD and other neurodegenerative diseases? As discussed previously, prion strains constitute an ensemble of PrP assemblies that are maintained under host selection, and drugs that target disease-associated species lead to the rapid development of resistance through strain adaptation<sup>6,37,91</sup>. However, targeting PrP<sup>C</sup>, which is the obligate substrate for all propagating or toxic species of PrP, is effective and can eliminate infection *in vivo* in preclinical studies with humanized anti-PrP<sup>C</sup> monoclonal antibodies (A. Khalili-Shirazi *et al.*, unpublished observations) (Fig. 2c). The propagation of prions is thought to involve the recruitment of largely unfolded PrP (ref. 89), and ligands that bind to the folded domain of PrP can act as a pharmacological chaperone that shifts the equilibrium in favour of the native state and inhibits propagation<sup>90</sup>. But effective targeting of an aggregation-prone, unstructured amyloid- $\beta$  peptide monomer with a ligand might be more challenging. Much investment in therapeutic agents for AD has focused on developing monoclonal antibodies that target amyloid- $\beta$  assemblies, with limited or no evidence of clinical benefits so far<sup>135</sup>. Because there is a considerable level of structural polymorphism in fibrillar amyloid- $\beta$ <sup>125,136</sup>, a concern — by analogy with prion disease — is that treatment with a specific monoclonal antibody might result in the targeting of only a subset of the ensemble of propagating species and lead to the selection of resistant seeds. In this context, a polyclonal response might be required to effectively target the molecular ensemble. Alternatively, APP expression could be targeted but such a strategy in established AD would require that blocking amyloid- $\beta$ -related neurotoxicity was still worthwhile. If the propagation of tau assemblies are an important factor in neurotoxicity, the effective treatment of established AD would require a dual therapy that targets tau as well as amyloid- $\beta$ . More generic approaches such as enhancing clearance mechanisms for misfolded proteins or pathological protein assemblies — restoring proteostasis — to reduce their propagation and spread are conceivable<sup>137</sup> and could be a valuable adjunct to highly targeted therapies.

### Perspective on future work

The long-term focus on prion disease has led to important insights into fundamental molecular processes that could be of wider relevance to pathobiology. The prion concept has been extended successfully from mammalian neurodegenerative diseases to encompass protein-based inheritance in yeast and other fungi, and the field of yeast prions has produced considerable advances in our understanding of prion strains and their structure<sup>2</sup>. However, with the growing consensus that mammalian and yeast prions are composed of fibrillar assemblies or amyloids, it is important to recognize that high-resolution structures of infectious prions and the structural basis of strain diversity are yet to be established. Amyloids can seed and propagate under suitable circumstances, and systemic amyloids may, at least in experiments, pass from host to host<sup>138</sup>. However, it is clear that amyloid fibrils can be produced from PrP that are not transmissible<sup>69</sup>. Prions appear to have a more complex quaternary structure. They may be better considered as a subset of fibrillar assemblies with structural characteristics that render them able to propagate *in vivo* at a level exceeding both the clearance mechanisms of the host and the dilution effects of cell division. Consequently, they are able to stably colonize a host, and the adaptability of the prion strain ensemble might be essential to such effective colonization. Such infection may or may not necessarily cause disease; however, subclinical infections are well recognized<sup>8,50–53</sup>. So, although prions might have the tinctorial and structural hallmarks of amyloids (Fig. 1), not all amyloids can act as prions.



The main goals of future work are to obtain high-resolution structures of infectious mammalian prions and to understand the structural basis of prion strain diversity. Such structural insights, which should be facilitated by technical advances in cryo-electron microscopy, would enable the rational, structure-based definition of prions. In turn, this would improve understanding of why some assemblies of misfolded proteins can act as efficient pathogens whereas others are benign. It will also be important to isolate and structurally characterize PrP<sup>L</sup>, as well as to understand how this neurotoxic species is generated. These advances could provide insights into the pathogenetic roles of the amyloid- $\beta$ , tau and  $\alpha$ -synuclein assemblies of other neurodegenerative diseases, which might act as propagating seeds or neurotoxic species. The development and validation of analytical tools — notably, robust assays for relevant seeding activity — will be needed to determine whether similar mechanistic relationships between propagating and neurotoxic species of prions are relevant to AD and other neurodegenerative diseases.

Achieving an accurate classification system for human prion strains that uses defined molecular criteria is also an important goal. Structure-based classification will enable aetiological and epidemiological studies to evaluate comprehensively the connection between prion strains that cause disease in humans with those that cause disease in animals. The knowledge acquired from understanding these relationships will have direct applications in the protection of public health and may lead to the reclassification of sporadic CJD, some cases of which could be due to unrecognized environmental prion exposure.

Prion diseases offer considerable advantages for research into therapeutic agents. First, laboratory animals are naturally susceptible to prion diseases, showing a complete and highly consistent clinico-pathological progression, which provides greater confidence that findings from animal models can be translated to humans. Second, CJD is a progressive and invariably fatal disorder with an average clinical duration of only four months. When combined with genetic stratification and new clinical tools, this enables the rapid assessment of drug efficacy from the study of only a few dozen patients<sup>139</sup>. Also, the levels of prions or disease-related PrPs can be measured directly in the blood or cerebrospinal fluid as a therapeutic biomarker that is more akin to measuring viral load in HIV trials than following neurodegenerative biomarkers of uncertain relationship to disease activity<sup>140,141</sup>. This should facilitate the speedy assessment of candidate therapeutic agents in a way that has been unfeasible so far with AD. Cellular PrP has been validated as a therapeutic target and passive immunization with anti-PrP monoclonal antibodies has provided a preclinical proof of principle, although it will be challenging to achieve therapeutic concentrations in the brain. Peripheral treatment with such antibodies, if safe, should enable the eradication of prions in asymptomatic infected individuals before neuroinvasion occurs. Estimates of the prevalence of vCJD-related prion infection in the UK population suggests that 1 in 2,000 people might be affected and, in addition, several thousand people have been exposed to prion-contaminated blood or blood products<sup>142</sup>. Experimental medicine studies, which might include the intracerebroventricular administration of antibodies in CJD, should provide important insights into the capacity of the brain to recover from the dementing process.

Evidence to support the iatrogenic transmission of amyloid- $\beta$  pathology has arisen necessarily from observational — rather than experimental — studies. It will be important to examine amyloid- $\beta$  seeding activity in archived batches of c-hGH and dura mater in suitable animal models and, with appropriate consent and ethical approval, to longitudinally study a subset of recipients of potentially contaminated batches of c-hGH who are at risk of developing iCJD and amyloid- $\beta$  pathology. Such studies would help to understand the risk posed to this group. They would also provide considerable insights into the development of amyloid- $\beta$  pathology, CAA and AD by following the evolution of cerebrospinal fluid biomarkers, serial amyloid, tau and vascular neuroimaging and neuropsychological and clinical features in a group of people who have been inoculated with amyloid- $\beta$  seeds during a defined period of time. The investigation of the temporal relationships between

amyloid- $\beta$ , tau deposition and neurodegeneration, as well as the study of early intervention with amyloid- $\beta$ -modifying therapeutic agents, should then be possible.

The transmission of amyloid- $\beta$  or other proteopathic seeds may be a rare event that is associated with discontinued medical practices; however, it could also turn out to be a wider problem. The use of blood products and tissue and organ grafting is a central part of modern medicine. Amyloid- $\beta$  seeds, similarly to prions, are known to adhere avidly to metal surfaces and to resist conventional sterilization<sup>143</sup>, and it will be important to consider whether they too can be transmitted by surgical instruments, blood transfusion, blood products or tissue grafting. Although it remains to be established whether iatrogenic exposure to amyloid- $\beta$  seeds can lead to AD, it seems clear that such exposure can lead to CAA that, in addition to cerebral haemorrhage, may also result in dementia.

Epidemiological studies until now have not raised concerns but new studies to investigate these specific issues are needed. Causal relationships will be more difficult to establish than for iCJD given the high prevalence of AD and related pathologies, as well as the frequency of surgical procedures, in an ageing population. The likelihood of incubation periods that span decades suggests it is prudent to consider a precautionary approach and to develop methods for removing protein seeds from surgical or medical instruments. Such methods were developed to remove prions from surgical instruments but were not used widely, and public-health authorities will need to balance the costs and benefits of developing these procedures in a situation for which the risks are uncertain<sup>144</sup>. ■

Received 8 June; accepted 9 September 2016.

- Collinge, J. Prion diseases of humans and animals: their causes and molecular basis. *Annu. Rev. Neurosci.* **24**, 519–550 (2001).
- Wickner, R. B. *et al.* Prion diseases of yeast: amyloid structure and biology. *Semin. Cell Dev. Biol.* **22**, 469–475 (2011).
- Jucker, M. & Walker, L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45–51 (2013).
- Bruce, M. E. Scrapie strain variation and mutation. *Br. Med. Bull.* **49**, 822–838 (1993).
- Collinge, J. Prion strain mutation and selection. *Science* **328**, 1111–1112 (2010).
- Li, J., Browning, S., Mahal, S. P., Oelschlegel, A. M. & Weissmann, C. Darwinian evolution of prions in cell culture. *Science* **327**, 869–872 (2010).
- An important study of prions in cell culture showing that biologically ‘cloned’ populations of prions gradually become heterogeneous by accumulating ‘mutants’, with selection pressure resulting in the emergence of different mutants in the evolving population.**
- Sandberg, M. K., Al Doujaily, H., Sharps, B., Clarke, A. R. & Collinge, J. Prion propagation and toxicity *in vivo* occur in two distinct mechanistic phases. *Nature* **470**, 540–542 (2011).
- Refs 7 and 19 demonstrate that prion propagation and neurotoxicity occur in two distinct mechanistic phases *in vivo*; they also show that neurotoxicity relates to distinct species of PrP that are produced following a pathway switch that occurs when prion levels become saturated.**
- Hill, A. F. *et al.* Species barrier independent prion replication in apparently resistant species. *Proc. Natl Acad. Sci. USA* **97**, 10248–10253 (2000).
- Griffith, J. S. Self Replication and scrapie. *Nature* **215**, 1043–1044 (1967).
- Alper, T., Cramp, W. A., Haig, D. A. & Clarke, M. C. Does the agent of scrapie replicate without nucleic acid? *Nature* **214**, 764–766 (1967).
- Prusiner, S. B. Prions. *Proc. Natl Acad. Sci. USA* **95**, 13363–13383 (1998).
- Telling, G. C. *et al.* Prion propagation in mice expressing human and chimeric PrP transgenes implicates the interaction of cellular PrP with another protein. *Cell* **83**, 79–90 (1995).
- Bolton, D. C., McKinley, M. P. & Prusiner, S. B. Identification of a protein that purifies with the scrapie prion. *Science* **218**, 1309–1311 (1982).
- Meyer, R. K. *et al.* Separation and properties of cellular and scrapie prion proteins. *Proc. Natl Acad. Sci. USA* **83**, 2310–2314 (1986).
- Gajdusek, D. C. Transmissible and non-transmissible amyloidoses: autocatalytic post-translational conversion of host precursor proteins to  $\beta$ -pleated sheet configurations. *J. Neuroimmunol.* **20**, 95–110 (1988).
- Come, J. H., Fraser, P. E. & Lansbury, P. T. J. A kinetic model for amyloid formation in the prion diseases: importance of seeding. *Proc. Natl Acad. Sci. USA* **90**, 5959–5963 (1993).
- Safar, J. *et al.* Eight prion strains have PrP<sup>Sc</sup> molecules with different conformations. *Nature Med.* **4**, 1157–1165 (1998).
- Cronier, S. *et al.* Detection and characterization of proteinase K-sensitive disease-related prion protein with thermolysin. *Biochem. J.* **416**, 297–305 (2008).
- Sandberg, M. K. *et al.* Prion neuropathology follows the accumulation of

- alternate prion protein isoforms after infective titre has peaked. *Nature Commun.* **5**, 4347 (2014).
20. Fraser, H. & Dickinson, A. G. Scrapie in mice: agent-strain differences in the distribution and intensity of grey matter vacuolation. *J. Comp. Pathol.* **83**, 29–40 (1973).
21. Bessen, R. A. & Marsh, R. F. Distinct PrP properties suggest the molecular basis of strain variation in transmissible mink encephalopathy. *J. Virol.* **68**, 7859–7868 (1994).
22. Collinge, J., Sidle, K. C., Meads, J., Ironside, J. & Hill, A. F. Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD. *Nature* **383**, 685–690 (1996).
23. Telling, G. C. *et al.* Evidence for the conformation of the pathologic isoform of the prion protein enciphering and propagating prion diversity. *Science* **274**, 2079–2082 (1996).
24. Bessen, R. A. *et al.* Non-genetic propagation of strain-specific properties of scrapie prion protein. *Nature* **375**, 698–700 (1995).
25. Bruce, M. *et al.* Transmission of bovine spongiform encephalopathy and scrapie to mice: strain variation and the species barrier. *Phil. Trans. R. Soc. Lond. B* **343**, 405–411 (1994).
26. Prusiner, S. B. *et al.* Transgenic studies implicate interactions between homologous PrP isoforms in scrapie prion replication. *Cell* **63**, 673–686 (1990).
27. Collinge, J. *et al.* Unaltered susceptibility to BSE in transgenic mice expressing human prion protein. *Nature* **378**, 779–783 (1995).
28. Collinge, J. Variant Creutzfeldt–Jakob disease. *Lancet* **354**, 317–323 (1999).
29. Collinge, J. & Clarke, A. A general model of prion strains and their pathogenicity. *Science* **318**, 930–936 (2007).
30. Wadsworth, J. D. *et al.* Human prion protein with valine 129 prevents expression of variant CJD phenotype. *Science* **306**, 1793–1796 (2004).
31. Hill, A. F. *et al.* The same prion strain causes vCJD and BSE. *Nature* **389**, 448–450 (1997).
32. Bruce, M. E. & Dickinson, A. G. Biological evidence that scrapie agent has an independent genome. *J. Gen. Virol.* **68**, 79–89 (1987).
33. Kimberlin, R. H. & Walker, C. A. Evidence that the transmission of one source of scrapie agent to hamsters involves separation of agent strains from a mixture. *J. Gen. Virol.* **39**, 487–496 (1978).
34. Polymenidou, M. *et al.* Coexistence of multiple PrP<sup>Sc</sup> types in individuals with Creutzfeldt–Jakob disease. *Lancet Neurol.* **4**, 805–814 (2005).
35. Yull, H. M. *et al.* Detection of type 1 prion protein in variant Creutzfeldt–Jakob disease. *Am. J. Pathol.* **168**, 151–157 (2006).
36. Taylor, D. M., Fernie, K., McConnell, I. & Steele, P. J. Observations on thermostable subpopulations of the unconventional agents that cause transmissible degenerative encephalopathies. *Vet. Microbiol.* **64**, 33–38 (1998).
37. Oelschlegel, A. M. & Weissmann, C. Acquisition of drug resistance and dependence by prions. *PLoS Pathog.* **9**, e1003158 (2013).
- Refs 37 and 91 demonstrate that prion populations or quasiespecies can develop drug resistance rapidly through the selection of resistant conformers.**
38. Sforzosi, L. *et al.* Cross-linking cellular prion protein triggers neuronal apoptosis *in vivo*. *Science* **303**, 1514–1516 (2004).
39. Hegde, R. S. *et al.* A transmembrane form of the prion protein in neurodegenerative disease. *Science* **279**, 827–834 (1998).
40. Ma, J., Wollmann, R. & Lindquist, S. Neurotoxicity and neurodegeneration when PrP accumulates in the cytosol. *Science* **298**, 1781–1785 (2002).
41. Chesebro, B. *et al.* Anchorless prion protein results in infectious amyloid disease without clinical scrapie. *Science* **308**, 1435–1439 (2005).
42. Sonati, T. *et al.* The toxicity of antiprion antibodies is mediated by the flexible tail of the prion protein. *Nature* **501**, 102–106 (2013).
43. Büeler, H. *et al.* Normal development and behaviour of mice lacking the neuronal cell-surface PrP protein. *Nature* **356**, 577–582 (1992).
44. Mallucci, G. R. *et al.* Post-natal knockout of prion protein alters hippocampal CA1 properties, but does not result in neurodegeneration. *EMBO J.* **21**, 202–210 (2002).
45. Büeler, H. *et al.* Mice devoid of PrP are resistant to scrapie. *Cell* **73**, 1339–1347 (1993).
46. Manson, J. C., Clarke, A., McBride, P. A., McConnell, I. & Hope, J. PrP gene dosage determines the timing but not the final intensity or distribution of lesions in scrapie pathology. *Neurodegeneration* **3**, 331–340 (1994).
47. Brandner, S. *et al.* Normal host prion protein necessary for scrapie-induced neurotoxicity. *Nature* **379**, 339–343 (1996).
48. Mallucci, G. *et al.* Depleting neuronal PrP in prion infection prevents disease and reverses spongiosis. *Science* **302**, 871–874 (2003).
49. Collinge, J. *et al.* Kuru in the 21st century—an acquired human prion disease with very long incubation periods. *Lancet* **367**, 2068–2074 (2006).
50. Race, R., Raines, A., Raymond, G. J., Caughey, B. & Chesebro, B. Long-term subclinical carrier state precedes scrapie replication and adaptation in a resistant species: analogies to bovine spongiform encephalopathy and variant Creutzfeldt–Jakob disease in humans. *J. Virol.* **75**, 10106–10112 (2001).
51. Hill, A. F. & Collinge, J. Subclinical prion infection. *Trends Microbiol.* **11**, 578–584 (2003).
52. Thackray, A. M., Klein, M. A., Aguzzi, A. & Bujdosó, R. Chronic subclinical prion disease induced by low-dose inoculum. *J. Virol.* **76**, 2510–2517 (2002).
53. Thackray, A. M., Klein, M. A. & Bujdosó, R. Subclinical prion disease induced by oral inoculation. *J. Virol.* **77**, 7991–7998 (2003).
54. Goedert, M. Alzheimer's and Parkinson's diseases: the prion concept in relation to assembled A $\beta$ , tau, and  $\alpha$ -synuclein. *Science* **349**, 1255555 (2015).
55. Walsh, D. M. & Selkoe, D. J. A critical appraisal of the pathogenic protein spread hypothesis of neurodegeneration. *Nature Rev. Neurosci.* **17**, 251–260 (2016).
- A counterreview to the extensive literature on prion-like mechanisms in neurodegeneration, which questions the role of propagating protein assemblies in pathogenesis and emphasises the role of selective neuronal vulnerability.**
56. Fraser, H. & Dickinson, A. G. Targeting of scrapie lesions and spread of agent via the retino-tectal projection. *Brain Res.* **346**, 32–41 (1985).
57. Vickery, C. M., Beck, K. E., Simmons, M. M., Hawkins, S. A. & Spiropoulos, J. Disease characteristics of bovine spongiform encephalopathy following inoculation into mice via three different routes. *Int. J. Exp. Pathol.* **94**, 320–328 (2013).
58. Bruce, M. E., Fraser, H., McBride, P. A., Scott, J. R. & Dickinson, A. G. *In Prion Diseases of Humans and Animals* (eds Prusiner, S. B., Collinge, J., Powell, J. & Anderton, B.) Ch. 40 (Ellis Horwood, 1992).
59. Schmidt, C. *et al.* A systematic investigation of production of synthetic prions from recombinant prion protein. *Open Biol.* **5**, 150165 (2015).
60. Salvadores, N., Shahnawaz, M., Scarpini, E., Tagliavini, F. & Soto, C. Detection of misfolded A $\beta$  oligomers for sensitive biochemical diagnosis of Alzheimer's disease. *Cell Rep.* **7**, 261–268 (2014).
- A proof-of-principle adaptation of the protein misfolding cyclic amplification method to prions to amplify amyloid- $\beta$  oligomers, which suggests the possibility of quantitative *in vitro* analysis of amyloid- $\beta$  seeding activity.**
61. Herva, M. E. *et al.* Anti-amyloid compounds inhibit  $\alpha$ -synuclein aggregation induced by protein misfolding cyclic amplification (PMCA). *J. Biol. Chem.* **289**, 11897–11905 (2014).
62. Diaz-Espinoza, R. & Soto, C. High-resolution structure of infectious prion protein: the final frontier. *Nature Struct. Mol. Biol.* **19**, 370–377 (2012).
63. Requena, J. R. & Wille, H. The structure of the infectious prion protein: experimental data and molecular models. *Prion* **8**, 60–66 (2014).
64. Jackson, G. S. *et al.* Reversible conversion of monomeric human prion protein between native and fibrillogenic conformations. *Science* **283**, 1935–1937 (1999).
65. Hornemann, S. & Glockshuber, R. A scrapie-like unfolding intermediate of the prion protein domain PrP(121–231) induced by acidic pH. *Proc. Natl Acad. Sci. USA* **95**, 6010–6014 (1998).
66. Daude, N., Lehmann, S. & Harris, D. A. Identification of intermediate steps in the conversion of a mutant prion protein to a scrapie-like form in cultured cells. *J. Biol. Chem.* **272**, 11604–11612 (1997).
67. Morillas, M., Vanik, D. L. & Surewicz, W. K. On the mechanism of  $\alpha$ -helix to  $\beta$ -sheet transition in the recombinant prion protein. *Biochemistry* **40**, 6982–6987 (2001).
68. Torrent, J. *et al.* High pressure induces scrapie-like prion protein misfolding and amyloid fibril formation. *Biochemistry* **43**, 7162–7170 (2004).
69. Tattum, M. H. *et al.* Elongated oligomers assemble into mammalian PrP amyloid fibrils. *J. Mol. Biol.* **357**, 975–985 (2006).
70. Collinge, J. *et al.* Transmission of fatal familial insomnia to laboratory animals. *Lancet* **346**, 569–570 (1995).
71. Lasmezás, C. I. *et al.* Transmission of the BSE agent to mice in the absence of detectable abnormal prion protein. *Science* **275**, 402–405 (1997).
72. Legname, G. *et al.* Synthetic mammalian prions. *Science* **305**, 673–676 (2004).
73. Deleault, N. R., Harris, B. T., Rees, J. R. & Supattapone, S. Formation of native prions from minimal components *in vitro*. *Proc. Natl Acad. Sci. USA* **104**, 9741–9746 (2007).
74. Makarava, N. *et al.* Recombinant prion protein induces a new transmissible prion disease in wild-type animals. *Acta Neuropathol.* **119**, 177 (2010).
75. Kim, J. I. *et al.* Mammalian prions generated from bacterially expressed prion protein in the absence of any mammalian cofactors. *J. Biol. Chem.* **285**, 14083–14087 (2010).
76. Wang, F., Wang, X., Yuan, C. G. & Ma, J. Generating a prion with bacterially expressed recombinant prion protein. *Science* **327**, 1132–1135 (2010).
77. Edgeworth, J. A. *et al.* Spontaneous generation of mammalian prions. *Proc. Natl Acad. Sci. USA* **107**, 14402–14406 (2010).
78. Prusiner, S. B. *et al.* Scrapie prions aggregate to form amyloid-like birefringent rods. *Cell* **35**, 349–358 (1983).
79. Safar, J. G. *et al.* Search for a prion-specific nucleic acid. *J. Virol.* **79**, 10796–10806 (2005).
80. Silveira, J. R. *et al.* The most infectious prion protein particles. *Nature* **437**, 257–261 (2005).
81. Caughey, B. & Lansbury, P. T., Jr. Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu. Rev. Neurosci.* **26**, 267–298 (2003).
82. Kraus, A., Groveman, B. R. & Caughey, B. Prions and the potential transmissibility of protein misfolding diseases. *Annu. Rev. Microbiol.* **67**, 543–564 (2013).
83. McKinley, M. P. *et al.* Scrapie prion rod formation *in vitro* requires both detergent extraction and limited proteolysis. *J. Virol.* **65**, 1340–1351 (1991).
84. Wenborn, A. *et al.* A novel and rapid method for obtaining high titre intact prion strains from mammalian brain. *Sci. Rep.* **5**, 10062 (2015).
85. Terry, C. *et al.* *Ex vivo* mammalian prions are formed of paired double helical prion protein fibrils. *Open Biol.* **6**, 160035 (2016).
- Precise cell-culture-based prion infectivity assays were used to define the physical relationship between PrP rods and prion infectivity, and electron tomography was used to define their architecture.**
86. Khalili-Shirazi, A. *et al.* PrP glycoforms are associated in a strain-specific ratio in native PrP<sup>Sc</sup>. *J. Gen. Virol.* **86**, 2635–2644 (2005).
87. Manson, J. C. *et al.* 129/Ola mice carrying a null mutation in PrP that abolishes mRNA production are developmentally normal. *Mol. Neurobiol.* **8**, 121–127 (1994).



88. Mallucci, G. R. *et al.* Targeting cellular prion protein reverses early cognitive deficits and neurophysiological dysfunction in prion-infected mice. *Neuron* **53**, 325–335 (2007).
89. Hosszu, L. P. *et al.* Structural mobility of the human prion protein probed by backbone hydrogen exchange. *Nature Struct. Biol.* **6**, 740–743 (1999).
90. Nicoll, A. J. *et al.* Pharmacological chaperone for the structured domain of human prion protein. *Proc. Natl Acad. Sci. USA* **107**, 17610–17615 (2010).
91. Berry, D. B. *et al.* Drug resistance confounding prion therapeutics. *Proc. Natl Acad. Sci. USA* **110**, E4160–E4169 (2013).
92. Enari, M., Flechsig, E. & Weissmann, C. Scrapie prion protein accumulation by scrapie-infected neuroblastoma cells abrogated by exposure to a prion protein antibody. *Proc. Natl Acad. Sci. USA* **98**, 9295–9299 (2001).
93. Peretz, D. *et al.* Antibodies inhibit prion propagation and clear cell cultures of prion infectivity. *Nature* **412**, 739–743 (2001).
94. Antonyuk, S. V. *et al.* Crystal structure of human prion protein bound to a therapeutic antibody. *Proc. Natl Acad. Sci. USA* **106**, 2554–2558 (2009).
95. White, A. R. *et al.* Monoclonal antibodies inhibit prion replication and delay the development of prion disease. *Nature* **422**, 80–83 (2003).
96. Song, C. H. *et al.* Effect of intraventricular infusion of anti-prion protein monoclonal antibodies on disease progression in prion-infected mice. *J. Gen. Virol.* **89**, 1533–1544 (2008).
97. Ohsawa, N. *et al.* Therapeutic effect of peripheral administration of an anti-prion protein antibody on mice infected with prions. *Microbiol. Immunol.* **57**, 288–297 (2013).
98. Klyubin, I. *et al.* Peripheral administration of a humanized anti-PrP antibody blocks Alzheimer's disease A $\beta$  synaptotoxicity. *J. Neurosci.* **34**, 6140–6145 (2014).
99. Gajdusek, D. C., Gibbs, C. J. Jr & Alpers, M. P. Experimental transmission of a kuru-like syndrome to chimpanzees. *Nature* **209**, 794–796 (1966).
100. Gibbs, C. J. Jr. *et al.* Creutzfeldt–Jakob disease (spongiform encephalopathy): transmission to the chimpanzee. *Science* **161**, 388–389 (1968).
101. Brown, P. *et al.* Human spongiform encephalopathy: the National Institutes of Health series of 300 cases of experimentally transmitted disease. *Ann. Neurol.* **35**, 513–529 (1994).
102. Baker, H. F., Ridley, R. M., Duchon, L. W., Crow, T. J. & Bruton, C. J. Induction of  $\beta$ (A4)-amyloid in primates by injection of Alzheimer's disease brain homogenate: comparison with transmission of spongiform encephalopathy. *Mol. Neurobiol.* **8**, 25–39 (1994).
- Refs 102 and 103 are important studies that demonstrate the seeding of amyloid- $\beta$  pathology in the primate brain by intracerebral inoculation with tissue affected by AD.**
103. Ridley, R. M., Baker, H. F., Windle, C. P., & Cummings, R. M. Very long term studies of the seeding of  $\beta$ -amyloidosis in primates. *J. Neural Transm.* **113**, 1243–1251 (2006).
104. Gandy, S. Lifelong management of amyloid- $\beta$  metabolism to prevent Alzheimer's disease. *N. Engl. J. Med.* **367**, 864–866 (2012).
105. Benilova, I., Karran, E. & De Strooper, B. The toxic A $\beta$  oligomer and Alzheimer's disease: an emperor in need of clothes. *Nature Neurosci.* **15**, 349–357 (2012).
106. Walsh, D. M. & Teplow, D. B. Alzheimer's disease and the amyloid  $\beta$ -protein. *Prog. Mol. Biol. Transl. Sci.* **107**, 101–124 (2012).
107. Jan, A., Hartley, D. M. & Lashuel, H. A. Preparation and characterization of toxic A $\beta$  aggregates for structural and functional studies in Alzheimer's disease research. *Nature Protoc.* **5**, 1186–1209 (2010).
108. Kane, M. D. *et al.* Evidence for seeding of  $\beta$ -amyloid by intracerebral infusion of Alzheimer brain extracts in  $\beta$ -amyloid precursor protein-transgenic mice. *J. Neurosci.* **20**, 3606–3611 (2000).
109. Meyer-Luehmann, M. *et al.* Exogenous induction of cerebral  $\beta$ -amyloidogenesis is governed by agent and host. *Science* **313**, 1781–1784 (2006).
110. Stöhr, J. *et al.* Purified and synthetic Alzheimer's amyloid beta (A $\beta$ ) prions. *Proc. Natl Acad. Sci. USA* **109**, 11025–11030 (2012).
111. Morales, R., Duran-Aniotz, C., Castilla, J., Estrada, L. D. & Soto, C. *De novo* induction of amyloid- $\beta$  deposition *in vivo*. *Mol. Psychiatry* **17**, 1347 (2012).
112. Eisele, Y. S. *et al.* Peripherally applied A $\beta$ -containing inoculates induce cerebral  $\beta$ -amyloidosis. *Science* **330**, 980–982 (2010).
- Provides evidence to show that amyloid- $\beta$  seeds can be transported from the abdomen to induce amyloid- $\beta$  deposition in the CNS in a mouse model.**
113. Clavaguera, F. *et al.* Peripheral administration of tau aggregates triggers intracerebral tauopathy in transgenic mice. *Acta Neuropathol.* **127**, 299–301 (2014).
114. Prusiner, S. B. *et al.* Evidence for  $\alpha$ -synuclein prions causing multiple system atrophy in humans with parkinsonism. *Proc. Natl Acad. Sci. USA* **112**, E5308–E5317 (2015).
115. Ward, H. J. *et al.* Sporadic Creutzfeldt–Jakob disease and surgery: a case-control study using community controls. *Neurology* **59**, 543–548 (2002).
116. Mahillo-Fernandez, I. *et al.* Surgery and risk of sporadic Creutzfeldt–Jakob disease in Denmark and Sweden: registry-based case-control studies. *Neuroepidemiology* **31**, 229–240 (2008).
117. Jaunmuktane, Z. *et al.* Evidence for human transmission of amyloid- $\beta$  pathology and cerebral amyloid angiopathy. *Nature* **525**, 247–250 (2015).
- A report of the possible iatrogenic transmission of amyloid- $\beta$  pathology and CAA many years after treatment with pituitary extracts derived from cadavers.**
118. Irwin, D. J. *et al.* Evaluation of potential infectivity of Alzheimer and Parkinson disease proteins in recipients of cadaver-derived human growth hormone. *JAMA Neurol.* **70**, 462–468 (2013).
119. Frontzek, K., Lutz, M. I., Aguzzi, A., Kovacs, G. G. & Budka, H. Amyloid- $\beta$  pathology and cerebral amyloid angiopathy are frequent in iatrogenic Creutzfeldt–Jakob disease after dural grafting. *Swiss Med. Wkly.* **146**, w14287 (2016).
- Refs 119 and 120 provide evidence that amyloid- $\beta$  pathology and CAA might also be iatrogenically transmitted to humans through neurosurgical procedures involving dura mater grafts.**
120. Kovacs, G. G. *et al.* Dura mater is a potential source of A $\beta$  seeds. *Acta Neuropathol.* **131**, 911–923 (2016).
121. Brandner, S. *et al.* Central and peripheral pathology of kuru: pathological analysis of a recent case and comparison with other forms of human prion disease. *Phil. Trans. R. Soc. B* **363**, 3755–3763 (2008).
122. Wadsworth, J. D. *et al.* Tissue distribution of protease resistant prion protein in variant CJD using a highly sensitive immuno-blotting assay. *Lancet* **358**, 171–180 (2001).
123. Urwin, P. J., Mackenzie, J. M., Llewelyn, C. A., Will, R. G. & Hewitt, P. E. Creutzfeldt–Jakob disease and blood transfusion: updated results of the UK Transfusion Medicine Epidemiology Review Study. *Vox Sang.* **110**, 310–316 (2016).
124. Haley, N. J. & Hoover, E. A. Chronic wasting disease of cervids: current knowledge and future perspectives. *Annu. Rev. Anim. Biosci.* **3**, 305–325 (2014).
125. Petkova, A. T. *et al.* Self-propagating, molecular-level polymorphism in Alzheimer's  $\beta$ -amyloid fibrils. *Science* **307**, 262–265 (2005).
126. Clavaguera, F. *et al.* Brain homogenates from human tauopathies induce tau inclusions in mouse brain. *Proc. Natl Acad. Sci. USA* **110**, 9535 (2013).
127. Guo, J. L. *et al.* Distinct  $\alpha$ -synuclein strains differentially promote tau inclusions in neurons. *Cell* **154**, 103–117 (2013).
128. Bousset, L. *et al.* Structural and functional characterization of two  $\alpha$ -synuclein strains. *Nature Commun.* **4**, 2575 (2013).
129. Heilbronner, G. *et al.* Seeded strain-like transmission of  $\beta$ -amyloid morphotypes in APP transgenic mice. *EMBO Rep.* **14**, 1017–1022 (2013).
130. Lu, J. X. *et al.* Molecular structure of  $\beta$ -amyloid fibrils in Alzheimer's disease brain tissue. *Cell* **154**, 1257–1268 (2013).
- A molecular structural model for amyloid- $\beta$ 40 fibrils seeded from the brains of two people with AD, which suggests that fibrils in the brain might spread from a single site of nucleation and that structural variations in fibrils might correlate with variations in AD phenotype.**
131. Cohen, M. L. *et al.* Rapidly progressive Alzheimer's disease features distinct structures of amyloid- $\beta$ . *Brain* **138**, 1009–1022 (2015).
132. Watts, J. C. *et al.* Serial propagation of distinct strains of A $\beta$  prions from Alzheimer's disease patients. *Proc. Natl Acad. Sci. USA* **111**, 10323–10328 (2014).
133. Wadsworth, J. D., Asante, E. A. & Collinge, J. Contribution of transgenic models to understanding human prion disease. *Neuropathol. Appl. Neurobiol.* **36**, 576–597 (2010).
134. Asante, E. A. *et al.* Transmission properties of human PrP 102L prions challenge the relevance of mouse models of GSS. *PLoS Pathog.* **11**, e1004953 (2015).
135. Scheltens, P. *et al.* Alzheimer's disease. *Lancet* **388**, 10–6736 505–517 (2016).
136. Hatami, A., Albay, R. III, Monjazeb, S., Milton, S. & Glabe, C. Monoclonal antibodies against A $\beta$ 42 fibrils distinguish multiple aggregation state polymorphisms *in vitro* and in Alzheimer disease brain. *J. Biol. Chem.* **289**, 32131–32143 (2014).
137. Labbadia, J. & Morimoto, R. I. The biology of proteostasis in aging and disease. *Annu. Rev. Biochem.* **84**, 435–464 (2015).
138. Tjernberg, L. O., Rising, A., Johansson, J., Jaudszems, K. & Westermark, P. Transmissible amyloid. *J. Intern. Med.* **280**, 153–163 (2016).
139. Mead, S. *et al.* Clinical trial simulations based on genetic stratification and the natural history of a functional outcome measure in Creutzfeldt–Jakob disease. *JAMA Neurol.* **73**, 447–455 (2016).
140. Edgeworth, J. A. *et al.* Detection of prion infection in variant Creutzfeldt–Jakob disease: a blood-based assay. *Lancet* **377**, 487–493 (2011).
141. Sawyer, E. B., Edgeworth, J. A., Thomas, C., Collinge, J. & Jackson, G. S. Preclinical detection of infectivity and disease-specific PrP in blood throughout the incubation period of prion disease. *Sci Rep.* **5**, 17742 (2015).
142. Gill, O. N. *et al.* Prevalent abnormal prion protein in human appendixes after bovine spongiform encephalopathy epizootic: large scale survey. *Br. Med. J.* **347**, f5675 (2013).
- A study of more than 30,000 archived surgical appendix samples that searched for evidence of vCJD prion infection; it suggested that about 1 in 2,000 of the UK population might be infected.**
143. Eisele, Y. S. *et al.* Induction of cerebral  $\beta$ -amyloidosis: intracerebral versus systemic A $\beta$  inoculation. *Proc. Natl Acad. Sci. USA* **106**, 12926–12931 (2009).
144. House of Commons Science and Technology Committee. After the storm? UK blood safety and the risk of variant Creutzfeldt–Jakob disease <http://www.publications.parliament.uk/pa/cm201415/cmselect/cmsctech/327/327.pdf> (House of Commons Science and Technology Committee, 2014).

**Acknowledgements** I am grateful to R. Newton for the preparation of figures and to H. Saibil, D. Walsh, J. Wadsworth, G. Jackson and S. Mead for helpful comments on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The author declares competing financial interests: see [go.nature.com/2ekvuhb](http://go.nature.com/2ekvuhb). Readers are welcome to comment on the online version of this paper at [go.nature.com/2ekvuhb](http://go.nature.com/2ekvuhb). Correspondence should be addressed to J.C. (jc@prion.ucl.ac.uk).



# The activities of amyloids from a structural perspective

Roland Riek<sup>1</sup> & David S. Eisenberg<sup>2,3</sup>

**The aggregation of proteins into structures known as amyloids is observed in many neurodegenerative diseases, including Alzheimer's disease. Amyloids are composed of pairs of tightly interacting, many stranded and repetitive intermolecular  $\beta$ -sheets, which form the cross- $\beta$ -sheet structure. This structure enables amyloids to grow by recruitment of the same protein and its repetition can transform a weak biological activity into a potent one through cooperativity and avidity. Amyloids therefore have the potential to self-replicate and can adapt to the environment, yielding cell-to-cell transmissibility, prion infectivity and toxicity.**

**T**he term amyloid describes a particular type of elongated, unbranched protein fibril. It was introduced by physician Rudolf Virchow to describe a macroscopic tissue abnormality with a pale, waxy appearance that produced a positive result — characteristic of starch-like materials — in an iodine-staining reaction; notably, the Latin word for starch is *amylum*<sup>1</sup>. As defined by pathologists<sup>1</sup> who document amyloids in numerous diseases, the fibrils must be deposited in body tissues and fluoresce a green–yellow colour when stained with the dye Congo red and viewed between crossed polarizers. As defined by biophysical scientists, who also see amyloids as denatured protein aggregates<sup>2–4</sup> and associate them with microbial and cellular functions<sup>5</sup>, the fibrils must display the ‘cross- $\beta$ ’ diffraction pattern caused by the cross- $\beta$ -sheet motif when irradiated with X-rays (Box 1). The motif is composed of tightly interacting intermolecular  $\beta$ -sheets, and each  $\beta$ -sheet comprises thousands of identical copies of the same  $\beta$ -strand that stack through hydrogen bonding. The backbone amide hydrogen bonds maintain the  $\beta$ -strands at a spacing of 4.8 Å in the direction of the fibril. Two or more such  $\beta$ -sheets lie in parallel, 6–12 Å apart.

The study of amyloids over the past half century has revealed that numerous proteins form amyloid fibrils, some of which are functional and some of which are pathological<sup>6–11</sup>. More than 40 proteins are known to form pathogenic amyloid fibrils<sup>1</sup> and there are at least a dozen functional amyloid fibrils<sup>12,13</sup>, with new members of both classes being discovered constantly. Therefore, a fascinating feature of the amyloid state is that a single type of fibril is formed by a wide variety of proteins and is associated with a wide variety of functions — both beneficial and pathogenic. In this Review, our aim is to examine present knowledge of the structures of these fibrils and to establish a structure–activity relationship for amyloid diseases, with a focus on devastating neurodegenerative disorders. Crucial questions include: the extent to which amyloids are the causative entities in the disease process; the mechanisms by which amyloid diseases are transmitted from cell to cell; and the cause of amyloid-induced toxicity.

## The cross- $\beta$ -sheet structure at atomic resolution

Although the cross- $\beta$ -sheet structure was described 80 years ago and was found to be common to half a dozen different disease-related amyloid fibrils 19 years ago<sup>14</sup>, after which it was studied using cryo-electron microscopy (cryo-EM)<sup>15</sup>, important questions remained because no atomic-resolution structures were available. One question concerned the features of the cross- $\beta$ -sheet structure that account for the extreme

stability of the amyloid, given that  $\beta$ -sheets form and break up readily. Another question focused on why amyloid fibrils consist generally of a single protein and do not contain mixtures of proteins as would be expected if they were held together mainly by intermolecular backbone hydrogen bonds. Answers emerged from atomic-resolution X-ray structures of amyloid fibrils formed by short peptide segments of amyloid-forming proteins<sup>16–26</sup>. These structures were enabled by the discovery that such fibril-forming segments also form needle-shaped microcrystals in which the amyloid protofilaments span the entire length of the crystals<sup>27</sup>.

The atomic-resolution X-ray structures reveal that the basic cross- $\beta$ -sheet motif consists of a pair of tightly mating repetitive  $\beta$ -sheets (Fig. 1 and Box 1). When viewed along the axis of the protofilament, the two  $\beta$ -sheets adhere by the interdigitation of the side chains of the mating strands, much like the teeth of a zipper. For this reason, the dual-sheet motif is termed a ‘steric zipper’. The interface between the two sheets in almost all cross- $\beta$ -sheet structures is devoid of water. The intricate interlacing of the side chains shows that the formation of the fibrils depends on the sequence of the participating segments. That is, the segments that form these structures are self-complementary. The stability of the fibrils arises from several factors. One such factor is the hydrogen bonds that form between backbone amide groups that run up and down the  $\beta$ -sheets. Because each amide hydrogen bond is polar, lines of parallel hydrogen bonds point up and down the  $\beta$ -sheets. In this arrangement, the hydrogen bonds polarize one another, which creates a cooperative energy of formation<sup>28</sup>. Another factor is the van der Waals forces that form between the closely interacting pairs of  $\beta$ -sheets. A third factor is the increase in entropy of the water molecules that are released from the inner faces of the two  $\beta$ -sheets that meet so closely. And a further stabilizing factor is the interaction of side chains that run up and down the  $\beta$ -sheets. Interacting side chains include Tyr aromatic rings that stack because of  $\pi$ – $\pi$  interactions. Side chains of Asn, Gln, Thr and Ser form hydrogen bonds known as ladders, which run up and down the fibrils (for example, the Gln ladder in Fig. 1) (refs 16 and 17).

All amyloid-segment microcrystals show that the cross- $\beta$ -sheet motif is composed of two, almost infinite,  $\beta$ -sheets with a steric-zipper side-chain interface. However, the various cross- $\beta$ -sheet structures can be categorized according to several criteria: whether their  $\beta$ -strands are parallel or antiparallel; whether their  $\beta$ -sheets pack with the same surfaces (face-to-face packing) or different surfaces (face-to-back packing) adjacent to one another; and, whether the two closely packed  $\beta$ -sheets

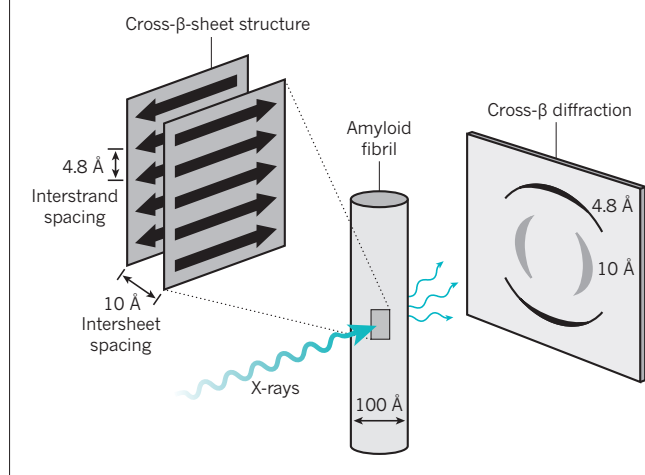
<sup>1</sup>Laboratory of Physical Chemistry, Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zürich, Switzerland. <sup>2</sup>UCLA-DOE Institute, Los Angeles, California 90095-1570, USA.

<sup>3</sup>Howard Hughes Medical Institute, Los Angeles, California 90095-1570, USA.

## BOX 1

The cross- $\beta$ -sheet motif

Amyloids are composed of an ordered, repetitive arrangement of many (usually thousands) copies of a peptide or protein (Box Fig., array of arrows). The repeating substructure consists of two layers of intermolecular  $\beta$ -sheets that run in the direction of the fibre axis. Amyloid fibrils can be identified easily using electron microscopy as long, unbranched filaments with diameters of 6–12 nm (ref. 129). The cross- $\beta$ -sheet motif gives rise to characteristic X-ray fibre diffraction patterns with a meridional reflection at about 4.8 Å, which corresponds to the spacing between  $\beta$ -strands, and a protein-dependent equatorial reflection at 6–12 Å, which corresponds to the distance between stacked  $\beta$ -sheets<sup>14,130</sup> (an example spacing of 10 Å is shown in the Box Fig.). The cross- $\beta$ -sheet motif was first described in 1935 by William Astbury, who measured the X-ray diffraction pattern of stretched, poached egg white<sup>130</sup>.



are oriented in the same direction (up–up) or opposite direction (up–down). Combinations of these three structural arrangements give eight theoretically possible classes of steric zippers, seven of which have now been seen in X-ray structures<sup>29</sup> (Fig. 1). Whether the  $\beta$ -strands in a  $\beta$ -sheet are parallel or antiparallel to each other, they are generally in register. This means that in a parallel  $\beta$ -sheet, each  $\beta$ -strand lies exactly above the  $\beta$ -strand below, and in an antiparallel  $\beta$ -sheet, each  $\beta$ -strand lies exactly above the  $\beta$ -strand that is two strands below. Out-of-register  $\beta$ -sheets have been seen in X-ray structures but are rare<sup>24,30</sup>.

The amyloid peptide crystal structures that we discuss indicate the prevalence of the cross- $\beta$ -sheet motif with steric-zipper-type side-chain interactions in peptide complementation and oligomerization in amyloid fibrils. However, an important consideration is the extent to which the structures of amyloid fibrils generated from segments of amyloid-forming proteins represent the fibril structures that are formed by their full-length parent proteins<sup>31</sup>.

### 3D structures of amyloid- $\beta$ , $\alpha$ -synuclein and HET-s prion

The relevance of the steric-zipper structures is supported by the observation of such a motif in a cryo-EM structure of amyloid- $\beta$ (1–42), which is a full-length amyloid fibril<sup>32</sup>. However, two — essentially identical — structures of amyloid- $\beta$ (1–42) fibrils determined by solid-state nuclear magnetic resonance (NMR) show that other types of interactions are also present<sup>33,34</sup> (Fig. 2). In these structures, there are also two peptides per layer of the protofilament axis, but each layer now contains two complete amyloid- $\beta$ (1–42) peptides. Although residues 1–14 of each peptide are poorly ordered, residues 15–42 form a double horseshoe shape, which resembles the letter S (ref. 35). Each

peptide contains four  $\beta$ -stands that stack on top of identical  $\beta$ -strands along the axis to form in-register, intermolecular  $\beta$ -sheets. Pairs of these strands interact with interdigitating or abutting side chains, but the pair members differ in sequence (Fig. 2b). Therefore, these zippers are heterosteric zippers, unlike the homosteric zippers that are found in the amyloid segments of the X-ray structures. Furthermore, each peptide has sharp bends at its Gly residues and contains two hydrophobic cores, an Asn ladder, a Gln ladder and a salt bridge between the side chain of Lys28 and the C terminus<sup>35</sup>, and its polar side chains face the solvent on the periphery of the protofilament. The two peptides of each layer meet at a two-fold axis (Fig. 2c) in a homosteric-zipper-like interaction; however, because each peptide contributes two segments to the zipper, it is a non-contiguous zipper interaction. Overall, the complexity of the fold seems to be dictated by the degree to which hydrophobic side chains can be buried, which also drives the folding of soluble proteins. Unusual structural features can arise from the energetic drive to bury hydrophobic side chains in structures that have not evolved to perform functions. Examples of these include sequential aromatic side chains (Phe19 and Phe20) that both face the hydrophobic core, and two negatively charged side chains (Glu22 and Asp23) in concert with Ala21 that face the solvent.

A 3D structure of  $\alpha$ -synuclein amyloid fibrils, which are pathogenic in neurons and are associated with Parkinson's disease (PD), has been determined by solid-state NMR<sup>36</sup> (Fig. 3a). The core structure (residues 46–94) has an orthogonal Greek-key topology. It is composed of several parallel, in-register cross- $\beta$ -sheets with heterosteric-zipper side-chain interactions such as hydrophobic interactions, aromatic  $\pi$ - $\pi$  stacking (Phe94) and glutamine ladders (for example, Gln79). It is flanked by less-structured N-terminal and C-terminal segments of more than 40 residues in length.

So far, no high-resolution structure of a mammalian prion in the amyloid state has been determined. A low-resolution electron tomography study of infective prion fibrils from the brain of a mouse has revealed paired protein fibrils<sup>37</sup> and a cryo-EM study of the infectious variant of prion protein (PrP<sup>Sc</sup>) from a mouse suggests that each molecule of PrP<sup>Sc</sup> has a  $\beta$ -solenoid structure with four layers<sup>38</sup>. By contrast, solid-state NMR experiments suggest that fibrils of recombinant PrP<sup>Sc</sup> have a parallel in-register, intermolecular  $\beta$ -sheet architecture<sup>39–42</sup>. In the absence of atomic-level information on mammalian prions, the high-resolution structure of the amyloid of infectious prion HET-s is noteworthy for its further elements of fibril architecture. HET-s is a functional prion from the filamentous fungus *Podospora anserina* and is involved in a primitive immune system<sup>43</sup>. The 3D structure of the infectious HET-s(218–289) fibrils determined by solid-state NMR reveals a left-handed  $\beta$ -solenoid that is composed of four in-register parallel  $\beta$ -sheets<sup>44</sup> (Fig. 3b). Each protein molecule contributes two windings of a helix to the length of the fibril so that the  $\beta$ -strands alternate between intermolecular and intramolecular hydrogen bonding along the axis of the fibre. The interior of the  $\beta$ -solenoid is composed of a hydrophobic core as well as Thr–Ser and Asn ladders, and the solvent-exposed face consists mostly of polar and charged side chains forming a pattern of alternating charges along the fibril axis that is enabled by the fibril's two-layer architecture. Notably, mammalian PrP<sup>Sc</sup> also seems to be composed of a  $\beta$ -solenoid but with four windings per molecule<sup>35,45</sup>.

### Structural polymorphism of amyloids

Amyloid polymorphism is a phenomenon in which a given peptide or protein sequence adopts two or more structurally distinct conformations of amyloid under the same environmental conditions. At the mesoscopic level, amyloid polymorphs can be identified using the features of their fibrils, including fibril morphologies (such as the ribbons or twisted fibrils of  $\alpha$ -synuclein)<sup>46,47</sup>, the degree of twist in the fibril, the number of protofilaments per fibril and the diameter or mass per unit of fibril length<sup>48</sup>, or by the detection of numerous signals per atom in solid-state NMR spectra<sup>47,49–52</sup>. At the atomic level, insights into the basis of the structural pluralism of amyloids can be obtained from

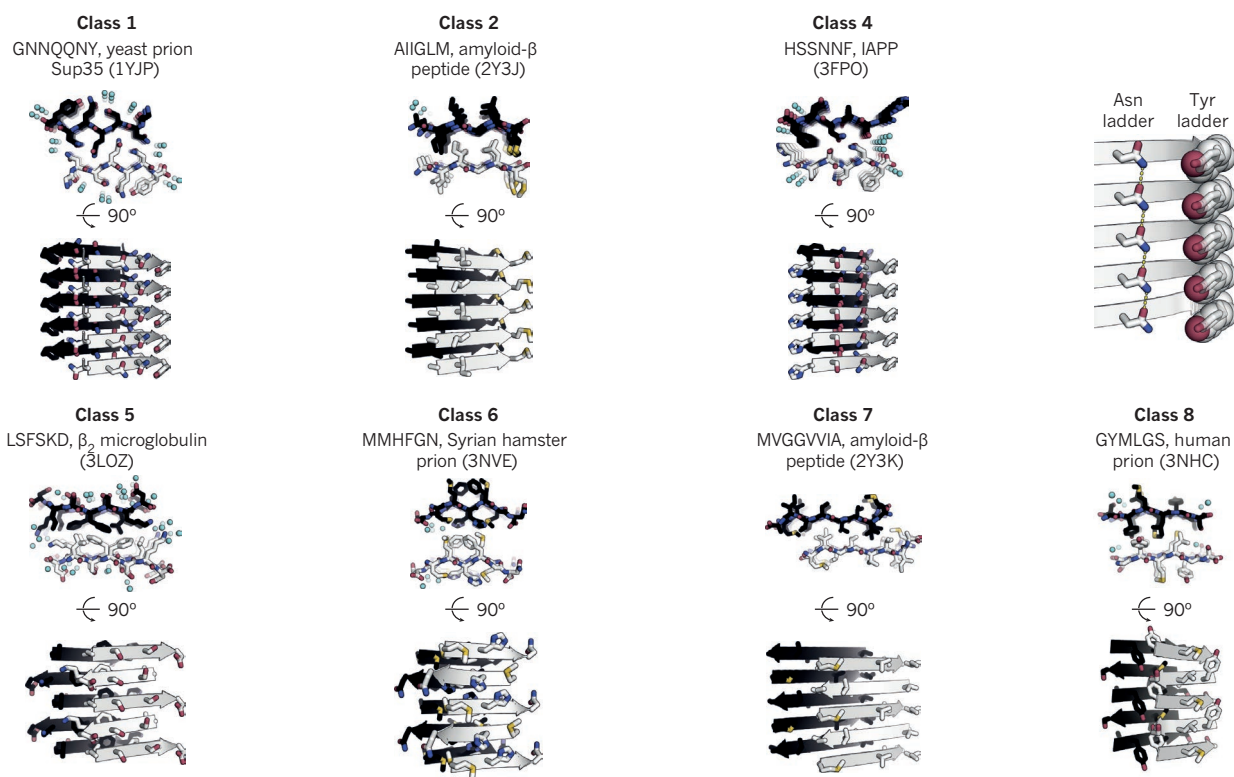
crystal structures of amyloid peptides (Fig. 4). For example, for the peptide segment KLVFAA (residues 16–21) of amyloid- $\beta$ (1–42), three steric zipper cross- $\beta$ -sheet motifs were found (Fig. 4a–c) and for the peptide segment MVGGVVIA (residues 35–42), two structures were found<sup>21</sup> (Fig. 4d, e). On the basis of these structures, such polymorphisms can be classified as segmental polymorphs in which different segments form the cross- $\beta$ -sheet cores of the two polymorphs, or as packing polymorphs in which the same segment forms the cross- $\beta$ -sheet core but is packed differently (such as parallel strands versus antiparallel strands or alternative steric-zipper side-chain interactions)<sup>20</sup>. At the mesoscopic level, polymorphisms can occur through different modes of binding between the protofilaments in a fibril. These weak, local interactions gain avidity by their summation over thousands of identical copies along the axis of the fibril. Such different supramolecular structures might be the result of underlying packing or segmental polymorphisms at the atomic level or could themselves be an independent form of polymorphism, termed an assembly polymorphism<sup>51</sup>. The number of amyloid polymorphs of a protein depends on the protein and can be estimated crudely as varying between a few and a few dozen<sup>9,50</sup>.

Amyloid polymorphism might seem to challenge Christian Anfinsen's thermodynamic hypothesis, which describes a one-to-one correlation between a polypeptide sequence and its 3D structure<sup>53</sup>. However, Anfinsen stated that his hypothesis applies to the structures of native proteins in their normal physiological milieu, in which the

native state is the lowest free-energy state under the given conditions. Pathogenic amyloid sequences are not in their native states when they form amyloid fibrils and, because they change conformation to reach the amyloid state, probably pass outside of their normal physiological milieu. Amyloid sequences that form pathogenic fibrils have not experienced evolutionary optimization, unlike functional amyloids<sup>54</sup>, which can yield energetically similar but distinct structures that are separated from one another by energy barriers larger than thermal energy. Furthermore, even a less stable polymorph can be the predominant species in a sample. This is because the aggregation of amyloid is initiated by stochastic nucleation and aggregates grow by kinetic processes in which the polymorph can replicate itself more quickly than a more stable polymorph. Together, these factors result in a plethora of kinetically uncoupled alternative structural states. For this reason, the collection of amyloid aggregates in an infected organism might comprise a pool of polymorphs with condition-dependent relative abundances. Because these polymorphs are structurally distinct, they could also have different activities. This is exemplified by the atomic-resolution structures shown in Fig. 4. The intersheet steric-zipper motifs between the polymorphs differ, yielding distinct amyloid stabilities, and correspondingly, the surfaces between the polymorphs require dissimilar interaction activities<sup>55</sup>.

## General properties of amyloids

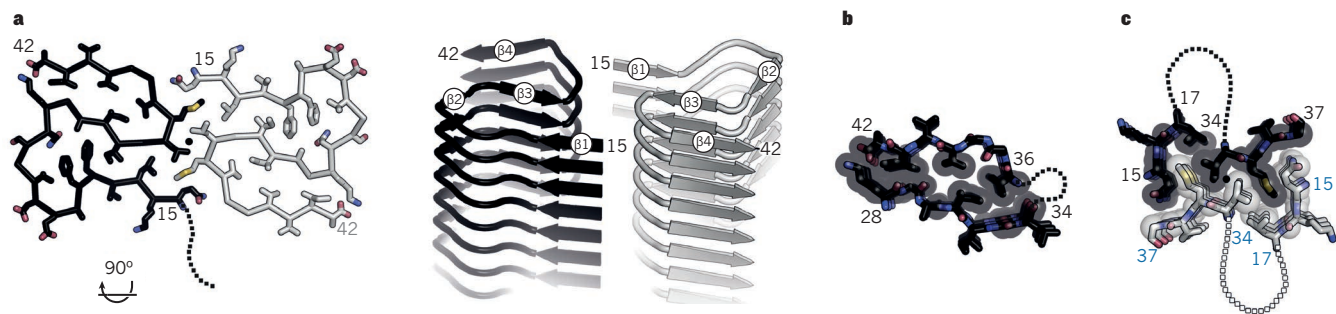
Two general biochemical properties account for the biological effects



**Figure 1 | Atomic-resolution X-ray diffraction steric-zipper structures of amyloid protofilaments formed by short segments of amyloid-forming proteins.** Each structure illustrates one of the possible symmetry classes of steric zippers, with views along (top) and perpendicular to (bottom) the protofilament axis. One  $\beta$ -sheet is shown in black and the other is in grey. Notably, water molecules (light blue spheres) are excluded from the tight interface between the  $\beta$ -sheets. Nitrogen atoms (blue), oxygen atoms (red), sulfur atoms (yellow) and  $\beta$ -strands (arrows) are depicted. For each class, the amino-acid sequence of the representative segment, the parent protein from which the segment was selected and the Protein Data Bank (PDB) accession code (in parentheses) for the coordinates of the structure is presented. Classes 1, 2 and 4 (top row) describe parallel  $\beta$ -sheets and classes 5–8 (bottom row) describe antiparallel  $\beta$ -sheets. Asn and Tyr ladders

are also shown (top right). Class 3 has not yet been observed experimentally. Each class is defined on the basis of several characteristics: whether their  $\beta$ -strands are parallel or antiparallel; whether their  $\beta$ -sheets pack with the same (face-to-face) surface or different (face-to-back) surface adjacent to one another; and, whether the two closely packed  $\beta$ -sheets are oriented in the same (up–up) direction or the opposite (up–down) direction. In summary: class 1 contains parallel, face-to-face, up–up structures; class 2 contains parallel, face-to-back, up–up structures; class 3 contains parallel, face-to-face, up–down structures; class 4 contains parallel, face-to-back, up–down structures; class 5 contains antiparallel, face-to-face, up–up structures. Class 6 contains antiparallel, face-to-back, up–up structures; class 7 contains antiparallel, face-to-face, up–down structures; and, class 8 contains antiparallel, face-to-face, up–down structures. IAPP, islet amyloid polypeptide.





**Figure 2 | Solid-state NMR structure of the ordered portion of the amyloid fibril formed by the AD-related peptide amyloid-β(1–42).** **a**, Views along the axis of the protofilament of amyloid-β(1–42) (left) and almost perpendicular to the same protofilament (right). Two essentially flat molecules in each layer of the protofilament are related by a two-fold axis (black circle). Each molecule has the shape of a double horseshoe, or an inverted letter S, and is stabilized by hydrogen bonding to molecules both

above and below, as well as by several heterosteric zippers. **b**, Example of a stabilizing heterosteric zipper. Note the tight packing of the side chains at the interface between the two segments of the zipper. **c**, The homosteric zipper that forms between the two molecules across the two-fold axis. Amino-acid residues from each molecule that contribute to the zipper interaction are not contiguous in the protein chain. In **b** and **c**, dotted lines represent intervening residues. Adapted with permission from ref. 128.

of amyloids. The first is that amyloid-forming proteins can switch from soluble monomers to insoluble fibrils (Box 2). This change of phase can disrupt the function of cells and organs, through either the loss of a crucial function or the gain of a toxic function. For example, the tumour-suppressor function of p53 is lost through p53 aggregation<sup>56–59</sup>. And the HET-s prion induces refolding of the soluble HET-S prion into a membrane protein that induces a necroptosis-related, spatially limited cell death in *Podospora anserina*<sup>60,61</sup>.

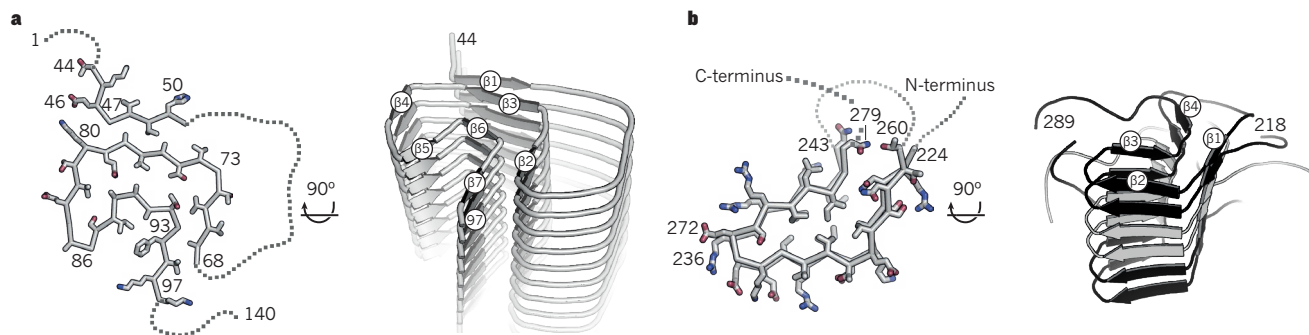
The second property is the repetitive structure of amyloids. Because the cross-β-sheet structure is built mostly of intermolecular contacts, the initial formation of amyloid is governed by a high concentration of amyloid protein, and after aggregation has been triggered (Box 2), amyloids might persist indefinitely (such as in the cases of the HET-s prion, which persists in nature<sup>60</sup>, and the amyloids that are proposed to be involved in memory<sup>62–64</sup>). Because of the subnanometre repetitiveness of the structure, a non-specific activity such as (inappropriate) ligand binding can be amplified into a potent effect through avidity or cooperativity. Its repetitive configuration also endows amyloids with a capacity to bind to other repetitive biomolecules such as RNA, DNA, glycosaminoglycans and lipid membranes with a relatively high affinity<sup>55,65,66</sup>. The structural repetitiveness of the amyloid fibre provides an ideal template for replication and might therefore be transmissible between cells — or even infectious in the case of prion diseases<sup>67</sup>.

### Replication and transmission of amyloids

The repetitive nature of amyloids is the basis of their capacity for templating identical structures through the recruitment of identical soluble protein molecules, therefore replicating themselves (Fig. 5). Replication

depends on the concentrations of soluble and aggregated amyloid proteins, the rate of growth of the amyloid and the rate of breakage of the amyloid fibre into seeds, and it might also depend on so-called secondary nucleation steps such as amyloid aggregation that is catalysed on the surface of the side of the fibril<sup>68–70</sup>. According to this theoretical framework, manipulation of these variables affects the growth and replication of amyloids<sup>71</sup>. For example, an increase in the levels of soluble amyloid protein in the host might be sufficient for the replication and spread of amyloid-β aggregates *in vivo* on inoculation with seed material (that is, amyloid templates)<sup>72</sup>, and the chaperone protein Hsp104, which has the ability to fragment fibrils of the Sup35 prion into smaller pieces, is required for prion propagation in yeast<sup>73</sup>.

Prions are the prototypical amyloids with *in vivo* replication properties. The ‘protein-only’ hypothesis<sup>74–76</sup> states that prion diseases, including scrapie in sheep, bovine spongiform encephalopathy (BSE) in cattle and Creutzfeldt–Jakob disease in humans, can be distinguished from infectious diseases that are caused by bacteria, viruses or viroids. Prion diseases originate and propagate through the conformational conversion of normal cellular prion protein (PrP<sup>C</sup>) into PrP<sup>Sc</sup> (ref. 67), an altered β-sheet-rich — and usually protease-resistant — form of PrP that is now known to be an amyloid entity. The amyloid serves as an infectious seed, which recruits soluble host PrP<sup>C</sup> for growth. The growing amyloid can fragment to produce further infectious seeds, which means that prions are essentially infectious proteins. Prions have been identified in lower eukaryotes, namely yeast and *Podospora anserina*<sup>43,77–80</sup>, as well as in mammals, and recombinant protein production has been used to show that the amyloid of a corresponding prion protein can be infectious in several of the kingdoms of life<sup>81–83</sup>.



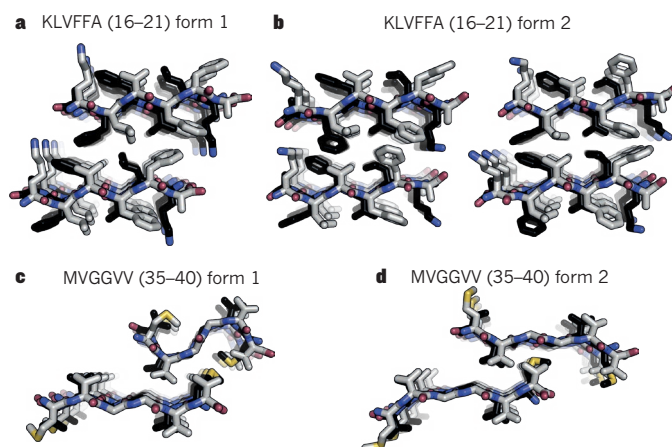
**Figure 3 | 3D structure of the cores of the α-synuclein fibrils and the HET-s(218–289) prion.** **a**, Views along the axis of the protofilament of α-synuclein (left) and almost perpendicular to the protofilament (right) (PDB accession code 2N0A). The overall architecture of the core of α-synuclein fibril comprising residues 44–97 is shown with nine molecules using a ribbon representation; β-strands are indicated (right). **b**, Views

along the axis of the protofilament of the HET-s(218–289) prion (left) and almost perpendicular to the protofilament (right). The overall architecture of the HET-s(218–289) prion fibril is shown with three molecules using a ribbon representation; the molecules are coloured black or grey and arrows indicate β-strands. Each molecule makes two windings in the β-solenoid structure.

Given the ability of amyloids to replicate, it is possible that amyloids other than prions might also spread from cell to cell *in vivo*, as well as from organism to organism. Cell-to-cell spreading of several disease-associated proteins in the amyloid state has been confirmed in laboratory animals<sup>84</sup>, but see Walsh and Selkoe<sup>85</sup>. Such amyloids include amyloid- $\beta$ <sup>72,86,87</sup>, which is associated with Alzheimer's disease (AD), and  $\alpha$ -synuclein<sup>88</sup>. The transmission of disease symptoms between mice has been shown for the amyloid-related condition amyloid A (AA) amyloidosis<sup>89</sup>. Furthermore, there is considerable evidence to suggest that the accumulation of amyloid- $\beta$  was induced in people through treatment with growth hormone prepared from large pools of cadaver-derived pituitary glands<sup>90</sup> and that brain homogenates from deceased people with AD induced the formation of plaques in marmosets<sup>91</sup>. The cellular spread of amyloid- $\beta$  is supported further by the presence of a single amyloid- $\beta$  polymorph in different regions of the brain in people with AD<sup>92</sup>. Although amyloid replication *in vivo* and its transmissibility are evident in many cases, it must be noted that, with the exception of prion diseases and systemic amyloidosis, the clinical symptoms of the diseases in these settings have been mostly missing<sup>93,94</sup>.

### Prion strains

In mammals, more than a dozen strains of prion are known<sup>95</sup>. Prion strains are distinguished by several phenotypes, including the length of incubation, the ability to cross a species barrier and the area of the brain in which PrP is deposited. Because PrP in its aggregated conformation is the infectious culprit of prion diseases (according to the protein-only hypothesis), the strain-specific properties must be attributed to altered conformations of PrP amyloids. On this basis, it is straightforward to propose that several distinct structural polymorphisms yield various forms of amyloid with distinct activities. Indeed, mammalian prion strains can be differentiated through patterns of PrP glycosylation<sup>96</sup>, conformation-selective PrP antibodies<sup>97</sup>, fluorescent dyes that bind to specific amyloid conformations<sup>98</sup> and PrP resistance to protease<sup>99</sup>. Furthermore, amyloids of the yeast prion Sup35 with distinct conformations can be propagated *in vitro* to produce several distinct strains of prion on infection<sup>68,100</sup>. Specifically, two polymorphs of Sup35 fibrils lead to different strains of prions with distinct cross- $\beta$ -sheet cores<sup>101</sup>.



**Figure 4 | Polymorphic atomic-resolution X-ray structures of amyloid fibrils formed by fragments of amyloid- $\beta$ , viewed along the axes of the fibrils.** **a, b**, Three polymorphic protofilaments of a segment of amyloid- $\beta$  with the sequence KLVFFA (residues 16–21). Actual fibrils have tens of thousands of layers but only four layers are shown. Alternating layers are coloured black and grey. **c, d**, Two polymorphic protofilaments of a fragment of amyloid- $\beta$  with the sequence MVGGVV (residues 35–40). Note that all  $\beta$ -sheets are antiparallel<sup>21</sup>.

### Adaptation of amyloids

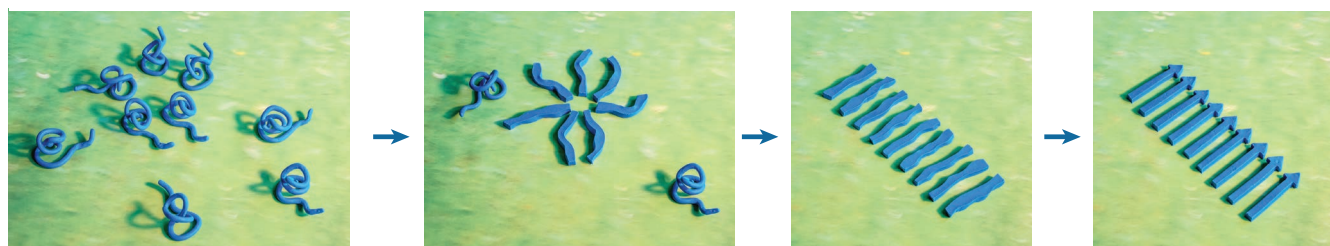
Strain adaptation is observed in prion diseases. For example, the adaptation of a prion strain to a new host results in a decrease in the time to the onset of symptoms for the first few serial transmissions. Several studies provide further evidence that prions are able to adapt to a new cellular environment<sup>102–104</sup>. In the presence of the drug swainsonine, a population of prions that were resistant to swainsonine emerged. Interestingly, after the drug was withdrawn, drug-sensitive prions appeared again. These findings indicate that the infectious material is composed of a pool of several prion polymorphisms, some of which are present only at very low levels. However, one such polymorphism might replicate more efficiently when the environment changes, becoming the predominant

## BOX 2

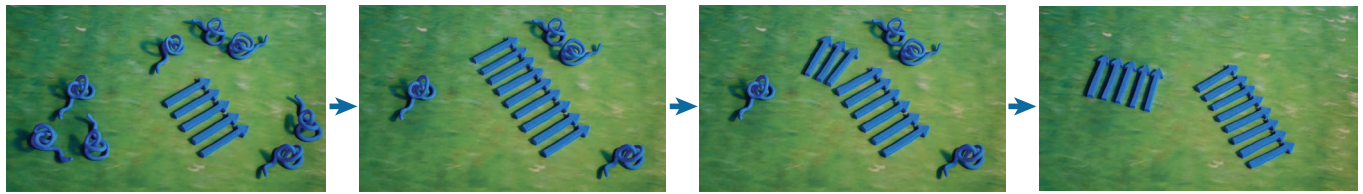
# Amyloid aggregation

A simplified mechanism of the nucleation-dependent process of amyloid aggregation<sup>75,118,131–133</sup> is shown in the Box Fig. For aggregation to be initiated, the soluble amyloid proteins or peptides (blue coils) must be partially unfolded, misfolded or intrinsically disordered. Furthermore, several molecules must come together to form a nucleus (star-like entity composed of several extended blue coils) from which the amyloid can grow. Because the nucleus is the most energetically unfavourable state, nucleus formation seems to determine the reaction rate. After the nucleus has formed, the further addition of monomers becomes faster, favouring growth of the aggregate (arrays of extended blue entities). Evidence from experiments indicates that such nuclei

could be oligomers of a different structural nature to that of the final amyloid fibrils (arrays of blue arrows). There might be further intermediates such as protofibrils (array of arrowless, extended blue entities) that finally become amyloid fibrils (array of blue arrows). Some fibril formation *in vitro* might also be driven by the artificial water–air interface<sup>134,135</sup> or could include helical intermediate states. Secondary nucleation processes, including fibril fragmentation and nucleus formation induced at the amyloid fibril surface<sup>69,70,136</sup>, are of particular relevance to the kinetics of fibril growth. Off-pathway aggregation might also occur. Notably, other mechanisms of aggregation such as downhill polymerization of transthyretin<sup>137</sup> can occur.







**Figure 5 | The replication of amyloids.** An amyloid seed (array of five blue arrows) recruits soluble amyloid protein (blue coils) and expands. The growing amyloid might fragment, which generates two seeds and therefore increases the number of replicative entities (arrays of five and eight arrows). The green background represents a given environment.

species<sup>105,106</sup> (Fig. 6). Even in the absence of a polymorph that is suited to the environment, a new polymorph could emerge through a *de novo* nucleation-dependent process of amyloid aggregation. Therefore, a species of amyloid protein has the potential to self-replicate after it has aggregated and can adapt to its environment, which suggests that it might evolve in a non-Mendelian fashion.

### Toxicity of amyloids

Although the pathological properties of amyloid load and propagation can be understood from the structural and physical-chemical properties of amyloids, the mechanism of amyloid toxicity in neurodegenerative diseases remains unknown. The association of amyloid fibrils with disease is strong: aggregates of amyloid- $\beta$  are found in plaques associated with AD; tau aggregates are found in neurofibrillary tangles that are also associated with AD;  $\alpha$ -synuclein fibrils are the main protein in Lewy bodies, the histological hallmark of PD<sup>107</sup>; and protease-resistant amyloid fibrils are found in prion diseases. Also, familial amyloid diseases are characterized by mutations that favour aggregation

by enhancing either the production of the amyloid protein<sup>108–110</sup> or by stabilizing the amyloid structure, as evident from the 3D structure of the amyloid- $\beta$ (1–42) amyloid (Fig. 2). All familial mutations in the amyloid- $\beta$ (1–42) sequence (that is, Lys16Asn, Ala21Gly, Glu22Gln, Glu22Gly, Glu22Lys, Glu22 $\Delta$  and Asp23Asn) are located at the structurally frustrated segment Ala20–Asp23, and most of these mutations are expected to release this frustration, at least partly. Although many of the mutations (Lys16Asn, Glu22Gln, Glu22Gly, Asp23Asn and Glu22Lys) probably attenuate intermolecular charge repulsion, Ala21Gly and Glu22Gly might instead relax the backbone restraints, owing to the peculiar side-chain arrangements of the segment Phe19–Asp23. Furthermore, in the rare familial neurodegenerative disease diabetes insipidus, amyloid aggregation of the prohormone provasopressin in the endoplasmic reticulum causes endoplasmic reticulum stress that results in cellular death<sup>111</sup>. Consequently, the amyloid state is associated strongly with disease. However, its association with direct cytotoxicity is less strong. The anchorless PrP<sup>Sc</sup> is highly infectious without causing toxicity<sup>112</sup>, and in other mammalian prion systems, the infectivity titre and toxicity seem to be uncoupled<sup>113</sup>. Furthermore, there seems to be a weak correlation between amyloid load and toxicity in AD<sup>114</sup>.

On the basis of these apparently inconsistent data, the following five hypotheses on the origin of toxicity have been proposed. First, amyloid fibrils are the toxic entities, as stated in the amyloid cascade hypothesis<sup>115–118</sup>. For example, on the basis of the high-resolution structure of amyloid- $\beta$ (1–42) fibrils, it is difficult to imagine that the presence of the hydrophobic repetitive patches of Val40–Ala42 flanked by Lys28 or Val18–Ala21 flanked by Arg16, which run along the surface of amyloid- $\beta$ (1–42) amyloids (Fig. 2) are without consequences for cell viability. Such regions might bind to other proteins tightly and without specificity through hydrophobic interactions or perturb or integrate, in part, with membranes.

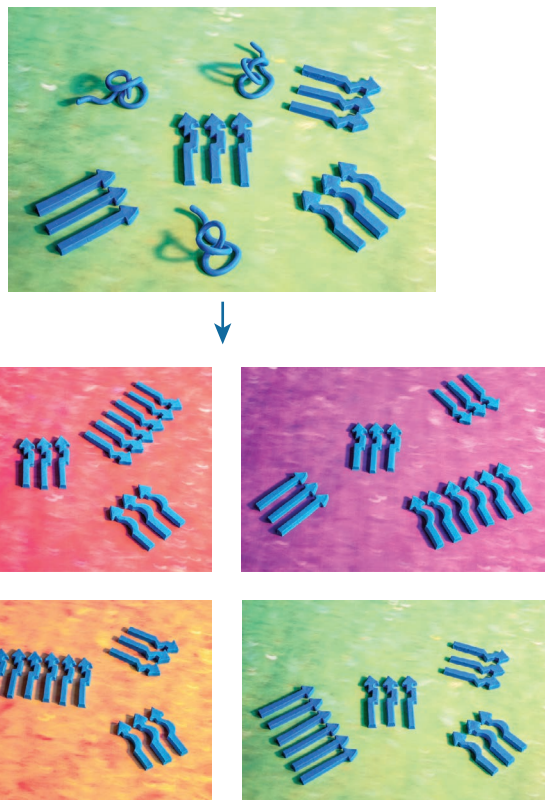
Second, amyloid oligomers are the toxic entities. These are smaller aggregates of the same proteins that go on to form fibrils<sup>114,119,120</sup> (Box 2). Amyloid fibrils might catalyse the formation of toxic oligomers through their surfaces<sup>69,121</sup> and could be the agents for cell transmission. This would make fibrils an important, although indirect, component of the mechanism of toxicity.

Third, other oligomers off the pathway to fibrils are the toxic entities. An example is cylindrin<sup>122</sup>, which might perturb the integrity of the cellular membrane by insertion. In this hypothesis, the amyloid fibril is a non-toxic entity that enables the cell to escape the toxicity caused by the oligomers.

Fourth, a process — rather than a specific structural state — accounts for the toxicity. That is, the mechanism of fibril growth *per se* is the toxic process<sup>123,124</sup>. This hypothesis is supported by the finding that during the aggregation of  $\alpha$ -synuclein on a lipid membrane, lipids are extracted by the growing fibril, whereas adding a fibril to a lipid bilayer merely results in binding of the amyloid to the membrane<sup>124</sup>.

Last, cellular stress accounts for the toxicity. The so-called neuronal vulnerability hypothesis states that neurodegenerative diseases spread in an organism: concomitantly, aggregation of the amyloid protein is a cellular stress that is transferred from cell to cell. The cellular stress not only degrades the cell but also induces *de novo* aggregation<sup>85</sup>.

When weighing up the various hypotheses, it must be noted that several amyloid polymorphs and various amyloid oligomers might be



**Figure 6 | The adaptation of amyloids to various environments.** Four different amyloid polymorphs (arrays of three polymorph-specific blue arrows, top) are surrounded by soluble amyloid protein (blue ribbons). When the environment is changed (red, pink or orange background) mainly one of the four polymorphs continues to grow, and each set of environmental conditions selects a different polymorph for expansion. With time, the growing polymorph becomes most abundant.



present in an affected individual and that each of these structural entities or mechanisms of aggregation could have distinct toxicity-related activities. The diseases might therefore be caused by several of these structural entities and processes, which would account for the complex, multifaceted nature of amyloid diseases that has been documented.

Given the limited fidelity of animal models to human diseases, tracing mechanisms of toxicity is especially challenging. Seeking chemical or biological interventions on the basis of structural and physical-chemical considerations might help to elucidate the structure-toxicity relationship in amyloid diseases. Possible approaches to intervention<sup>125</sup> include: the prevention of amyloid elongation through use of structure-based inhibitors (such as ref. 58); the inhibition of secondary nucleation<sup>121,126</sup>; the inhibition of amyloid nucleus or oligomer formation; the inhibition of amyloid oligomer membrane insertion; the inhibition of amyloid spreading between cells; and, the inhibition of amyloid formation through the stabilization of native structures. The structures of the amyloids might therefore be used in the rational design of molecules for disease intervention. For example, the hydrophobic hot spot around Phe19 and Phe20 in the structure of amyloid- $\beta$ (1–42) (Fig. 2) is a potential interesting pharmacophore for the inhibition of fibril elongation. The epitopes comprising residues Val40–Ala42–Lys28 or Leu17–Ala21–Arg16 form surface-exposed hydrophobic patches that are flanked by a positive charged ladder, indicating that they are involved in secondary nucleation or membrane interaction and making them prime candidates for intervention in these processes. Disease intervention through the stabilization of a native structure has yielded a drug for the condition transthyretin amyloidogenesis (reviewed in ref. 127). This physico-chemical approach, termed kinetic stabilization, is built on the hypothesis that the formation of amyloids causes disease and that amyloid aggregation can be inhibited by stabilizing the folded native tetramer of the protein transthyretin<sup>45</sup>. Remarkably, this approach enabled discovery of the drug tafamidis without the availability of good animal models of transthyretin amyloidogenesis. Tafamidis was recently approved by the European Medicines Agency for the treatment of transthyretin familial amyloid polyneuropathy. This example illustrates how a thorough understanding of the mechanism of aggregation of the soluble state of an amyloid protein can lead to a treatment for amyloid-based disease. Furthermore, it is a prime example of the establishment of a drug based purely on physical-chemical principles and studies and it is the first drug that has been approved to treat an amyloid disease. ■

Received 1 August; accepted 14 September 2016.

- Sipe, J. D. *et al.* Nomenclature 2014: amyloid fibril proteins and clinical classification of the amyloidosis. *Amyloid* **21**, 221–224 (2014).
- Gustavsson, A., Engström, U. & Westermark, P. Normal transthyretin and synthetic transthyretin fragments form amyloid-like fibrils *in vitro*. *Biochem. Biophys. Res. Commun.* **175**, 1159–1164 (1991).
- Colom, W. & Kelly, J. W. Partial denaturation of transthyretin is sufficient for amyloid fibril formation *in vitro*. *Biochemistry* **31**, 8654–8660 (1992).
- Booth, D. R. *et al.* Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* **385**, 787–793 (1997).
- Fowler, D. M., Koulou, A. V., Balch, W. E. & Kelly, J. W. Functional amyloid — from bacteria to humans. *Trends Biochem. Sci.* **32**, 217–224 (2007).
- Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).
- Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).
- Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nature Rev. Mol. Cell Biol.* **15**, 384–396 (2014).
- Greenwald, J. & Riek, R. Biology of amyloid: structure, function, and regulation. *Structure* **18**, 1244–1260 (2010).
- Eisenberg, D. & Jucker, M. The amyloid state of proteins in human diseases. *Cell* **148**, 1188–1203 (2012).
- Eisele, Y. S. *et al.* Targeting protein aggregation for the treatment of degenerative diseases. *Nature Rev. Drug Discov.* **14**, 759–780 (2015).
- Mauy, C. P. J. The emerging concept of functional amyloid. *J. Intern. Med.* **265**, 329–334 (2009).
- Maji, S. K. *et al.* Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* **325**, 328–332 (2009).
- Sunde, M. *et al.* Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **273**, 729–739 (1997).
- Evidence that amyloid fibrils, in general, are composed of the cross- $\beta$ -sheet structure.**
- Jiménez, J. L. *et al.* The protofibril structure of insulin amyloid fibrils. *Proc. Natl Acad. Sci. USA* **99**, 9196–9201 (2002).
- The first atomic-resolution structures of amyloid-like crystals, which revealed the steric-zipper side-chain motif, including the Asn and Gln hydrogen-bond ladders.**
- Nelson, R. *et al.* Structure of the cross- $\beta$  spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
- Sawaya, M. R. *et al.* Atomic structures of amyloid cross- $\beta$  spines reveal varied steric zippers. *Nature* **447**, 453–457 (2007).
- The various classes of cross- $\beta$ -sheet structures are revealed at atomic resolution.**
- Ivanova, M. I., Sievers, S. A., Sawaya, M. R., Wall, J. S. & Eisenberg, D. Molecular basis for insulin fibril assembly. *Proc. Natl Acad. Sci. USA* **106**, 18990–18995 (2009).
- Wiltzius, J. J. W., Sievers, S. A., Sawaya, M. R. & Eisenberg, D. Atomic structures of IAPP (amylin) fusions suggest a mechanism for fibrillation and the role of insulin in the process. *Protein Sci.* **18**, 1521–1530 (2009).
- Wiltzius, J. J. W. *et al.* Molecular mechanisms for protein-encoded inheritance. *Nature Struct. Mol. Biol.* **16**, 973–978 (2009).
- Colletier, J.-P. *et al.* Molecular basis for amyloid- $\beta$  polymorphism. *Proc. Natl Acad. Sci. USA* **108**, 16938–16943 (2011).
- Apostol, M. I., Sawaya, M. R., Cascio, D. & Eisenberg, D. Crystallographic studies of prion protein (PrP) segments suggest how structural changes encoded by polymorphism at residue 129 modulate susceptibility to human prion disease. *J. Biol. Chem.* **285**, 29671–29675 (2010).
- Apostol, M. I., Wiltzius, J. J. W., Sawaya, M. R., Cascio, D. & Eisenberg, D. Atomic structures suggest determinants of transmission barriers in mammalian prion disease. *Biochemistry* **50**, 2456–2463 (2011).
- Soriaga, A. B., Sangwan, S., Macdonald, R., Sawaya, M. R. & Eisenberg, D. Crystal structures of IAPP amyloidogenic segments reveal a novel packing motif of out-of-register beta sheets. *J. Phys. Chem. B* **120**, 5810–5816 (2016).
- Saelices, L. *et al.* Uncovering the mechanism of aggregation of human transthyretin. *J. Biol. Chem.* **290**, 28932–28943 (2015).
- Rodriguez, J. A. *et al.* Structure of the toxic core of  $\alpha$ -synuclein from invisible crystals. *Nature* **525**, 486–490 (2015).
- Balbirnie, M., Grothe, R. & Eisenberg, D. S. An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated  $\beta$ -sheet structure for amyloid. *Proc. Natl Acad. Sci. USA* **98**, 2375–2380 (2001).
- Tsemekhman, K., Goldschmidt, L., Eisenberg, D. & Baker, D. Cooperative hydrogen bonding in amyloid formation. *Protein Sci.* **16**, 761–764 (2007).
- Landreh, M. *et al.* The formation, function and regulation of amyloids: insights from structural biology. *J. Intern. Med.* **280**, 164–176 (2016).
- Liu, C. *et al.* Out-of-register  $\beta$ -sheets suggest a pathway to toxic amyloid aggregates. *Proc. Natl Acad. Sci. USA* **109**, 20913–20918 (2012).
- van der Wel, P. C. A., Lewandowski, J. R. & Griffin, R. G. Solid-state NMR study of amyloid nanocrystals and fibrils formed by the peptide GNNQQNY from yeast prion protein Sup35p. *J. Am. Chem. Soc.* **129**, 5117–5130 (2007).
- Schmidt, M. *et al.* Peptide dimer structure in an A $\beta$ (1–42) fibril visualized with cryo-EM. *Proc. Natl Acad. Sci. USA* **112**, 11858–11863 (2015).
- The cryo-EM reconstruction of an amyloid- $\beta$ (1–42) fibril reveals an intermolecular steric-zipper motif.**
- Colvin, M. T. *et al.* Atomic resolution structure of monomeric A $\beta$ 42 amyloid fibrils. *J. Am. Chem. Soc.* **138**, 9663–9674 (2016).
- Refs 33 and 34 present the atomic-resolution solid-state NMR structure of an amyloid- $\beta$ (1–42) fibril associated with AD, revealing hydrophobic interactions in concert with the cross- $\beta$ -sheet motif as important contributors to the amyloid structure.**
- Wälti, M. A. *et al.* Atomic-resolution structure of a disease-relevant A $\beta$ (1–42) amyloid fibril. *Proc. Natl Acad. Sci. USA* **113**, E4976–E4984 (2016).
- Xiao, Y. *et al.* A $\beta$ (1–42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nature Struct. Mol. Biol.* **22**, 499–505 (2015).
- Tuttle, M. D. *et al.* Solid-state NMR structure of a pathogenic fibril of full-length human  $\alpha$ -synuclein. *Nature Struct. Mol. Biol.* **23**, 409–415 (2016).
- A solid-state NMR structure of cell-infectious amyloid fibrils of  $\alpha$ -synuclein associated with PD.**
- Terry, C. *et al.* *Ex vivo* mammalian prions are formed of paired double helical prion protein fibrils. *Open Biol.* **6**, 160035 (2016).
- Vázquez-Fernández, E. *et al.* The structural architecture of an infectious mammalian prion using electron cryomicroscopy. *PLoS Pathog.* **12**, e1005835 (2016).
- Groveman, B. R. *et al.* Parallel inter-register intermolecular  $\beta$ -sheet architectures for prion-seeded prion protein (PrP) amyloids. *J. Biol. Chem.* **289**, 24129–24142 (2014).
- Helmus, J. J., Surewicz, K., Nadaud, P. S., Surewicz, W. K. & Jarosiec, C. P. Molecular conformation and dynamics of the Y145Stop variant of human prion protein in amyloid fibrils. *Proc. Natl Acad. Sci. USA* **105**, 6284–6289 (2008).
- Müller, H. *et al.* Progress towards structural understanding of infectious sheep PrP-amyloid. *Prion* **8**, 344–358 (2014).
- Jones, E. M. *et al.* Structural polymorphism in amyloids: new insights from studies with Y145Stop prion protein fibrils. *J. Biol. Chem.* **286**, 42777–42784 (2011).

43. Coustou, V., Deleu, C., Saupe, S. & Begueret, J. The protein product of the het-s heterokaryon incompatibility gene of the fungus *Podospora anserina* behaves as a prion analog. *Proc. Natl Acad. Sci. USA* **94**, 9773–9778 (1997).
44. Wasmer, C. et al. Amyloid fibrils of the HET-s(218–289) prion form a  $\beta$  solenoid with a triangular hydrophobic core. *Science* **319**, 1523–1526 (2008).  
**The 3D structure of a biological-relevant functional amyloid reveals a  $\beta$ -solenoid helix with two windings per molecules, indicating an evolutionarily-evolved prion.**
45. Klabunde, T. et al. Rational design of potent human transthyretin amyloid disease inhibitors. *Nature Struct. Biol.* **7**, 312–321 (2000).
46. Vilar, M. et al. The fold of  $\alpha$ -synuclein fibrils. *Proc. Natl Acad. Sci. USA* **105**, 8637–8642 (2008).
47. Bousset, L. et al. Structural and functional characterization of two  $\alpha$ -synuclein strains. *Nature Commun.* **4**, 2575 (2013).  
**Two  $\alpha$ -synuclein polymorphs with distinct structures at both the atomic and the mesoscopic levels show distinct toxicities.**
48. Meinhardt, J., Sachse, C., Hortschansky, P., Grigorieff, N. & Fändrich, M. A $\beta$ (1–40) fibril polymorphism implies diverse interaction patterns in amyloid fibrils. *J. Mol. Biol.* **386**, 869–877 (2009).
49. Ravotti, F. et al. Solid-state NMR sequential assignment of an amyloid- $\beta$ (1–42) fibril polymorph. *Biomol. NMR Assign.* **10**, 269–276 (2016).
50. Tycko, R. Physical and structural basis for polymorphism in amyloid fibrils. *Protein Sci.* **23**, 1528–1539 (2014).
51. Paravastu, A. K., Leapman, R. D., Yau, W.-M. & Tycko, R. Molecular structural basis for polymorphism in Alzheimer's  $\beta$ -amyloid fibrils. *Proc. Natl Acad. Sci. USA* **105**, 18349–18354 (2008).
52. Petkova, A. T. et al. Self-propagating, molecular-level polymorphism in Alzheimer's  $\beta$ -amyloid fibrils. *Science* **307**, 262–265 (2005).  
**The presence of self-propagating polymorphs is revealed at the atomic level using solid-state NMR.**
53. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
54. Daskalov, A. et al. Contribution of specific residues of the  $\beta$ -solenoid fold to HET-s prion function, amyloid structure and stability. *PLoS Pathog.* **10**, e1004158 (2014).
55. Wang, L., Schubert, D., Sawaya, M. R., Eisenberg, D. & Riek, R. Multidimensional structure-activity relationship of a protein in its aggregated states. *Angew. Chem. Int. Edn Engl.* **49**, 3904–3908 (2010).
56. Silva, J. L., De Moura Gallo, C. V., Costa, D. C. F. & Rangel, L. P. Prion-like aggregation of mutant p53 in cancer. *Trends Biochem. Sci.* **39**, 260–267 (2014).
57. Ano Bom, A. P. D. et al. Mutant p53 aggregates into prion-like amyloid oligomers and fibrils: implications for cancer. *J. Biol. Chem.* **287**, 28152–28162 (2012).
58. Soragni, A. et al. A designed inhibitor of p53 aggregation rescues p53 tumor suppression in ovarian carcinomas. *Cancer Cell* **29**, 90–103 (2016).
59. Xu, J. et al. Gain of function of mutant p53 by coaggregation with multiple tumor suppressors. *Nature Chem. Biol.* **7**, 285–295 (2011).
60. Riek, R. & Saupe, S. J. The HET-S/s prion motif in the control of programmed cell death. *Cold Spring Harb. Perspect. Biol.* **8**, a023515 (2016).
61. Seuring, C. et al. The mechanism of toxicity in HET-S/HET-s prion incompatibility. *PLoS Biol.* **10**, e1001451 (2012).
62. Caudron, F. & Barral, Y. A super-assembly of Whi3 encodes memory of deceptive encounters by single cells during yeast courtship. *Cell* **155**, 1244–1257 (2013).
63. Si, K., Choi, Y.-B., White-Grindley, E., Majumdar, A. & Kandel, E. R. Aplysia CPEB can form prion-like multimers in sensory neurons that contribute to long-term facilitation. *Cell* **140**, 421–435 (2010).
64. Si, K., Lindquist, S. & Kandel, E. R. A neuronal isoform of the aplysia CPEB has prion-like properties. *Cell* **115**, 879–891 (2003).
65. Geoghegan, J. C. et al. Selective incorporation of polyanionic molecules into hamster prions. *J. Biol. Chem.* **282**, 36341–36353 (2007).
66. Supattapone, S. Elucidating the role of cofactors in mammalian prion propagation. *Prion* **8**, 100–105 (2014).
67. Prusiner, S. B. Prions. *Proc. Natl Acad. Sci. USA* **95**, 13363–13383 (1998).
68. Tanaka, M., Chien, P., Naber, N., Cooke, R. & Weissman, J. S. Conformational variations in an infectious protein determine prion strain differences. *Nature* **428**, 323–328 (2004).
69. Cohen, S. I. A. et al. Proliferation of amyloid- $\beta$ 42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl Acad. Sci. USA* **110**, 9758–9763 (2013).
70. Arosio, P., Knowles, T. P. J. & Linse, S. On the lag phase in amyloid fibril formation. *Phys. Chem. Chem. Phys.* **17**, 7606–7618 (2015).
71. Riek, R. Cell biology: infectious Alzheimer's disease? *Nature* **444**, 429–431 (2006).
72. Meyer-Luehmann, M. et al. Exogenous induction of cerebral  $\beta$ -amyloidogenesis is governed by agent and host. *Science* **313**, 1781–1784 (2006).  
**Demonstrates that aggregates of amyloid- $\beta$ (1–42) are transmissible in a mouse model of AD.**
73. Chernoff, Y. O., Lindquist, S. L., Ono, B., Inge-Vechtomov, S. G. & Liebman, S. W. Role of the chaperone protein Hsp104 in propagation of the yeast prion-like factor [psi<sup>+</sup>]. *Science* **268**, 880–884 (1995).
74. Alper, T., Cramp, W. A., Haig, D. A. & Clarke, M. C. Does the agent of scrapie replicate without nucleic acid? *Nature* **214**, 764–766 (1967).
75. Griffith, J. S. Self-replication and scrapie. *Nature* **215**, 1043–1044 (1967).
76. Prusiner, S. B. Novel proteinaceous infectious particles cause scrapie. *Science* **216**, 136–144 (1982).
77. Serio, T. R. & Lindquist, S. L. [PSI<sup>+</sup>], SUP35, and chaperones. *Adv. Protein Chem.* **57**, 335–366 (2001).
78. Wickner, R. B. et al. Yeast prions act as genes composed of self-propagating protein amyloids. *Adv. Protein Chem.* **57**, 313–334 (2001).
79. Edsles, H. K., Gray, V. T. & Wickner, R. B. The [URE3] prion is an aggregated form of Ure2p that can be cured by overexpression of Ure2p fragments. *Proc. Natl Acad. Sci. USA* **96**, 1498–1503 (1999).
80. Patino, M. M., Liu, J. J., Glover, J. R. & Lindquist, S. Support for the prion hypothesis for inheritance of a phenotypic trait in yeast. *Science* **273**, 622–626 (1996).
81. Castilla, J., Saá, P., Hetz, C. & Soto, C. *In vitro* generation of infectious scrapie prions. *Cell* **121**, 195–206 (2005).
82. Wang, F., Wang, X., Yuan, C.-G. & Ma, J. Generating a prion with bacterially expressed recombinant prion protein. *Science* **327**, 1132–1135 (2010).
83. Ritter, C. et al. Correlation of structural elements and infectivity of the HET-s prion. *Nature* **435**, 844–848 (2005).
84. Jucker, M. & Walker, L. C. Pathogenic protein seeding in Alzheimer disease and other neurodegenerative disorders. *Ann. Neurol.* **70**, 532–540 (2011).
85. Walsh, D. M. & Selkoe, D. J. A critical appraisal of the pathogenic protein spread hypothesis of neurodegeneration. *Nature Rev. Neurosci.* **17**, 251–260 (2016).
86. Morales, R., Bravo-Alegria, J., Duran-Aniotz, C. & Soto, C. Titration of biologically active amyloid- $\beta$  seeds in a transgenic mouse model of Alzheimer's disease. *Sci. Rep.* **5**, 9349 (2015).
87. Wu, J. W. et al. Neuronal activity enhances tau propagation and tau pathology *in vivo*. *Nature Neurosci.* **19**, 1085–1092 (2016).
88. Luk, K. C. et al. Pathological  $\alpha$ -synuclein transmission initiates Parkinson-like neurodegeneration in nontransgenic mice. *Science* **338**, 949–953 (2012).
89. Lundmark, K. et al. Transmissibility of systemic amyloidosis by a prion-like mechanism. *Proc. Natl Acad. Sci. USA* **99**, 6979–6984 (2002).
90. Jaunmuktane, Z. et al. Evidence for human transmission of amyloid- $\beta$  pathology and cerebral amyloid angiopathy. *Nature* **525**, 247–250 (2015).  
**Demonstrates that amyloid- $\beta$ (1–42) pathology is transmissible in humans.**
91. Ridley, R. M., Baker, H. F., Windle, C. P. & Cummings, R. M. Very long term studies of the seeding of  $\beta$ -amyloidosis in primates. *J. Neural Transm.* **113**, 1243–1251 (2006).
92. Lu, J.-X. et al. Molecular structure of  $\beta$ -amyloid fibrils in Alzheimer's disease brain tissue. *Cell* **154**, 1257–1268 (2013).
93. Aguzzi, A. & Lakkaraju, A. K. K. Cell biology of prions and prionoids: a status report. *Trends Cell Biol.* **26**, 40–51 (2016).
94. Aguzzi, A. & Rajendran, L. The transcellular spread of cytosolic amyloids, prions, and prionoids. *Neuron* **64**, 783–790 (2009).
95. Weissmann, C. Birth of a prion: spontaneous generation revisited. *Cell* **122**, 165–168 (2005).
96. Asante, E. A. et al. BSE prions propagate as either variant CJD-like or sporadic CJD-like prion strains in transgenic mice expressing human prion protein. *EMBO J.* **21**, 6358–6366 (2002).
97. Cali, I. et al. Co-existence of scrapie prion protein types 1 and 2 in sporadic Creutzfeldt-Jakob disease: its effect on the phenotype and prion-type characteristics. *Brain* **132**, 2643–2658 (2009).
98. Sigurdson, C. J. et al. Prion strain discrimination using luminescent conjugated polymers. *Nature Methods* **4**, 1023–1030 (2007).
99. Cali, I. et al. Classification of sporadic Creutzfeldt-Jakob disease revisited. *Brain* **129**, 2266–2277 (2006).
100. King, C.-Y. & Diaz-Avalos, R. Protein-only transmission of three yeast prion strains. *Nature* **428**, 319–323 (2004).
101. Toyama, B. H., Kelly, M. J. S., Gross, J. D. & Weissman, J. S. The structural basis of yeast prion strain variants. *Nature* **449**, 233–237 (2007).
102. Li, J., Browning, S., Mahal, S. P., Oelschlegel, A. M. & Weissmann, C. Darwinian evolution of prions in cell culture. *Science* **327**, 869–872 (2010).  
**Evidence for the adaptation of prions.**
103. Weissmann, C., Li, J., Mahal, S. P. & Browning, S. Prions on the move. *EMBO Rep.* **12**, 1109–1117 (2011).
104. Oelschlegel, A. M. & Weissmann, C. Acquisition of drug resistance and dependence by prions. *PLoS Pathog.* **9**, e1003158 (2013).
105. Collinge, J. & Clarke, A. R. A general model of prion strains and their pathogenicity. *Science* **318**, 930–936 (2007).
106. Eigen, M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–523 (1971).
107. Spillantini, M. G. et al.  $\alpha$ -Synuclein in Lewy bodies. *Nature* **388**, 839–840 (1997).
108. Goldgaber, D., Lerman, M. I., McBride, O. W., Saffioti, U. & Gajdusek, D. C. Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease. *Science* **235**, 877–880 (1987).
109. Tanzi, R. E. et al. Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. *Science* **235**, 880–884 (1987).
110. Suzuki, N. et al. An increased percentage of long amyloid beta protein secreted by familial amyloid beta protein precursor (beta APP717) mutants. *Science* **264**, 1336–1340 (1994).
111. Birk, J., Friberg, M. A., Prescianotto-Baschong, C., Spiess, M. & Rutishauser, J. Dominant pro-vasopressin mutants that cause diabetes insipidus form disulfide-linked fibrillar aggregates in the endoplasmic reticulum. *J. Cell Sci.* **122**, 3994–4002 (2009).
112. Chesebro, B. et al. Anchorless prion protein results in infectious amyloid disease without clinical scrapie. *Science* **308**, 1435–1439 (2005).
113. Sandberg, M. K., Al-Doujaily, H., Sharps, B., Clarke, A. R. & Collinge, J. Prion propagation and toxicity *in vivo* occur in two distinct mechanistic phases. *Nature* **470**, 540–542 (2011).

114. Walsh, D. M. & Selkoe, D. J. A $\beta$  oligomers — a decade of discovery. *J. Neurochem.* **101**, 1172–1184 (2007).
  115. Hardy, J. A. & Higgins, G. A. Alzheimer's disease: the amyloid cascade hypothesis. *Science* **256**, 184–185 (1992).
  116. Silveira, J. R. *et al.* The most infectious prion protein particles. *Nature* **437**, 257–261 (2005).
  117. Pieri, L., Madiona, K., Bousset, L. & Melki, R. Fibrillar  $\alpha$ -synuclein and huntingtin exon 1 assemblies are toxic to the cells. *Biophys. J.* **102**, 2894–2905 (2012).
  118. Jarrett, J. T. & Lansbury, P. T. Seeding 'one-dimensional crystallization' of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **73**, 1055–1058 (1993).
  119. Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T. & Lansbury, P. T. Neurodegenerative disease: amyloid pores from pathogenic mutations. *Nature* **418**, 291 (2002).
  120. Abedini, A. *et al.* Time-resolved studies define the nature of toxic IAPP intermediates, providing insight for anti-amyloidosis therapeutics. *eLife* **5**, e12977 (2016).
  121. Cohen, S. I. A. *et al.* A molecular chaperone breaks the catalytic cycle that generates toxic A $\beta$  oligomers. *Nature Struct. Mol. Biol.* **22**, 207–213 (2015).
  122. Laganowsky, A. *et al.* Atomic view of a toxic amyloid small oligomer. *Science* **335**, 1228–1231 (2012).
  123. Jan, A. *et al.* A $\beta$ 42 neurotoxicity is mediated by ongoing nucleated polymerization process rather than by discrete A $\beta$ 42 species. *J. Biol. Chem.* **286**, 8585–8596 (2011).
  124. Reynolds, N. P. *et al.* Mechanism of membrane interaction and disruption by  $\alpha$ -synuclein. *J. Am. Chem. Soc.* **133**, 19366–19375 (2011).
  125. Ankarcona, M. *et al.* Current and future treatment of amyloid diseases. *J. Intern. Med.* **280**, 177–202 (2016).
  126. Knight, S. D., Presto, J., Linse, S. & Johansson, J. The BRICHOS domain, amyloid fibril formation, and their relationship. *Biochemistry* **52**, 7523–7531 (2013).
  127. Johnson, S. M., Connelly, S., Fearn, C., Powers, E. T. & Kelly, J. W. The transthyretin amyloidosis: from delineating the molecular mechanism of aggregation linked to pathology to a regulatory-agency-approved drug. *J. Mol. Biol.* **421**, 185–203 (2012).
- A review on the successful development of a drug through stabilizing the native form of the amyloid protein.**
128. Eisenberg, D. S. & Sawaya, M. R. Implications for Alzheimer's disease of an atomic resolution structure of amyloid- $\beta$ (1–42) fibrils. *Proc. Natl Acad. Sci. USA* **113**, 9398–9400 (2016).
  129. Sunde, M. & Blake, C. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Adv. Protein Chem.* **50**, 123–159 (1997).
  130. Astbury, W. T., Dickinson, S. & Bailey, K. The X-ray interpretation of denaturation and the structure of the seed globulins. *Biochem. J.* **29**, 2351–2360 (1935).
- The cross- $\beta$ -sheet structure is introduced on the basis of X-ray fibre diffraction studies of protein aggregates induced by heat denaturation.**
131. Asakura, S. Polymerization of flagellin and polymorphism of flagella. *Adv. Biophys.* **1**, 99–155 (1970).
  132. Oosawa, F. & Asakura, S. *Thermodynamics of the Polymerization of Protein* (Academic, 1975).
  133. Bishop, M. F. & Ferrone, F. A. Kinetics of nucleation-controlled polymerization. A perturbation treatment for use with a secondary pathway. *Biophys. J.* **46**, 631–644 (1984).
  134. Pronchik, J., He, X., Giurleo, J. T. & Talaga, D. S. *In vitro* formation of amyloid from  $\alpha$ -synuclein is dominated by reactions at hydrophobic interfaces. *J. Am. Chem. Soc.* **132**, 9797–9803 (2010).
  135. Campioni, S. *et al.* The presence of an air-water interface affects formation and elongation of  $\alpha$ -synuclein fibrils. *J. Am. Chem. Soc.* **136**, 2866–2875 (2014).
  136. Knowles, T. P. J. *et al.* An analytical solution to the kinetics of breakable filament assembly. *Science* **326**, 1533–1537 (2009).
  137. Hurshman, A. R., White, J. T., Powers, E. T. & Kelly, J. W. Transthyretin aggregation under partially denaturing conditions is a downhill polymerization. *Biochemistry* **43**, 7365–7381 (2004).

**Acknowledgements** We thank K. Corniotto and M. Sawaya for making figures and the Swiss National Science Foundation (SNSF), the US National Institutes of Health and the Howard Hughes Medical Institute for continuing support of our research, including an SNSF Sinergia grant to R.R.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this paper at [go.nature.com/2f8daq7](http://go.nature.com/2f8daq7). Correspondence should be addressed to R.R. ([roland.riek@phys.chem.ethz.ch](mailto:roland.riek@phys.chem.ethz.ch)) and D.S.E. ([david@mbi.ucla.edu](mailto:david@mbi.ucla.edu)).

**Reviewer Information** *Nature* thanks G. Pielak and the other anonymous reviewer(s) for their contribution to the peer review of this work.



# The stem osteichthyan *Andreolepis* and the origin of tooth replacement

Donglei Chen<sup>1</sup>, Henning Blom<sup>1</sup>, Sophie Sanchez<sup>2,3</sup>, Paul Tafforeau<sup>3</sup> & Per E. Ahlberg<sup>2</sup>

**The teeth of gnathostomes (jawed vertebrates) show rigidly patterned, unidirectional replacement that may or may not be associated with a shedding mechanism. These mechanisms, which are critical for the maintenance of the dentition, are incongruently distributed among extant gnathostomes. Although a permanent tooth-generating dental lamina is present in all chondrichthyans, many tetrapods and some teleosts, it is absent in the non-teleost actinopterygians. Tooth-shedding by basal hard tissue resorption occurs in most osteichthyans (including tetrapods) but not in chondrichthyans. Here we report a three-dimensional virtual dissection of the dentition of a 424-million-year-old stem osteichthyan, *Andreolepis hedei*, using propagation phase-contrast synchrotron microtomography, with a reconstruction of its growth history. *Andreolepis*, close to the common ancestor of all extant osteichthyans, shed its teeth by basal resorption but probably lacked a permanent dental lamina. This is the earliest documented instance of resorptive tooth shedding and may represent the primitive osteichthyan mode of tooth replacement.**

As humans, we tend to view teeth as a specific feature of the mouth. Yet tooth-like structures made of dentine are, in many fossil vertebrates and some living forms such as sharks, distributed over the body surface as well<sup>1,2</sup>. Teeth, denticles and other tooth-like dentine structures are collectively known as odontodes and are characterized by homologous tissues, developmental characteristics and gene-expression patterns<sup>3–5</sup>. Nevertheless, oral teeth differ from dermal denticles in both morphology and spatial arrangement. This spatial arrangement is, in part, a product of a tightly organized mode of tooth production<sup>4–10</sup>. In extant gnathostomes, new teeth can be produced either by a discrete epithelial fold known as the dental lamina or directly from the epithelium surrounding the standing tooth<sup>5–10</sup>. The dental lamina can be a small, temporary structure associated with an individual tooth or a large, permanent tissue fold at the lingual margin of the dentition. In osteichthyans, where the teeth are attached to dermal jaw bones, the dental lamina can burrow into the medullary cavity of this bone to produce intraosseous replacement teeth that rise from below (as in our own jaws) or be more superficial and produce extraosseous replacement teeth above the jaw bone<sup>6,8</sup>. Chondrichthyans and many tetrapods have a permanent dental lamina<sup>6</sup>, whereas actinopterygians typically produce replacement teeth next to individual standing teeth, with either a small successional (transient) dental lamina or none at all<sup>8–11</sup>. A shedding mechanism to dispose of old teeth is often associated with these tooth-generating mechanisms. In osteichthyans this involves the basal resorption of the old teeth<sup>12</sup>, as we see during the loss of our own milk teeth, when the roots are resorbed before the crowns are finally shed. In chondrichthyans, shedding is not accompanied by hard-tissue resorption; the entire tooth is shed, including the base.

The single, permanent dental lamina of extant chondrichthyans was originally interpreted as being primitive for gnathostomes and homologous with the dental lamina of tetrapods<sup>13,14</sup>. Data from an increasing range of actinopterygians and sarcopterygians (including tetrapods) paint a less straightforward picture, one that possibly involves multiple origins of the dental lamina<sup>7–11</sup>. To understand the early evolution of this system, data on tooth-replacement processes from the few taxa that can be assigned to the osteichthyan stem group, helping to polarize

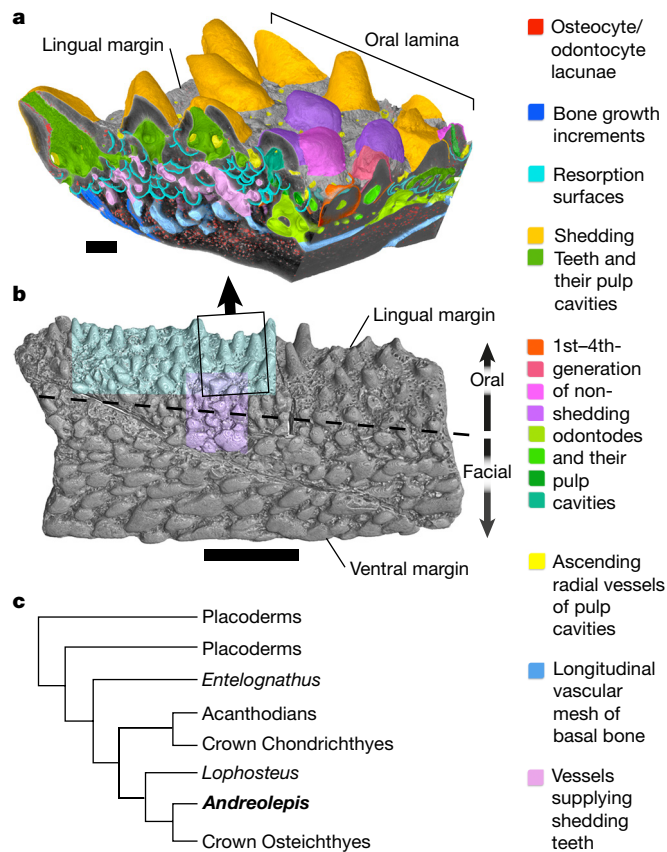
character differences among crown Osteichthyes<sup>1</sup>, are essential. Here we present such data from *Andreolepis hedei*.

## *Andreolepis*

*Andreolepis hedei* comes from the Late Silurian (Ludfordian) of Gotland, Sweden, and is approximately 424 million years old<sup>1,15–19</sup>. The material consists of disarticulated scales and dermal bone fragments, part of a diverse assemblage of microremains<sup>20</sup>. The scales of *Andreolepis*, of which there are hundreds of specimens, are characteristically osteichthyan both in morphology and in carrying odontodes covered with enamel (a unique osteichthyan tissue)<sup>1,17,18,21</sup>. Osteichthyan-like dermal bones from the same assemblage, including marginal jaw bones and small, cushion-shaped, tooth-bearing bones (but no tooth whorls) have been attributed to *Andreolepis*<sup>16,17</sup>, but differ from the scales in lacking enamel. However, the recent discovery that the early osteichthyan *Psarolepis* carries both enamel-covered and naked odontodes<sup>1</sup> shows that this difference does not invalidate the attribution. As there is only one type of osteichthyan scale present in the assemblage, we regard the attribution of the osteichthyan-like jaw bones to *Andreolepis* as secure. The lack of enamel on the teeth and facial odontodes, together with the absence of lepidotrichia (characteristic of crown-group osteichthyan fishes) in the assemblage<sup>17</sup>, supports the identification of *Andreolepis* as a stem osteichthyan<sup>1,22</sup> (Fig. 1).

The original interpretation of specimen LOG87-301DF, probably a right dentary with a biting surface covered in small teeth, on the basis of external observation was that it carried non-shedding, obliquely transverse tooth files<sup>17</sup>. However, a subsequent study of a similar specimen, NRM-PZ P.15910, using virtual thin sections produced by synchrotron microtomography, dismissed this 'dentition' as dermal ornament and questioned the identification of these elements as jaws<sup>23</sup>. The exhaustive investigation into specimen LOG87-301DF that we present here, using data from propagation phase-contrast synchrotron microtomography (PPC-SRμCT) scans made at beamline ID19 of the European Synchrotron Radiation Facility (ESRF), Grenoble, France, reaches completely different conclusions from the NRM-PZ P.15910 study<sup>23</sup> and allows us to reinterpret the scan data from the latter. This reflects two key factors. First, that the higher resolution of our data reveals

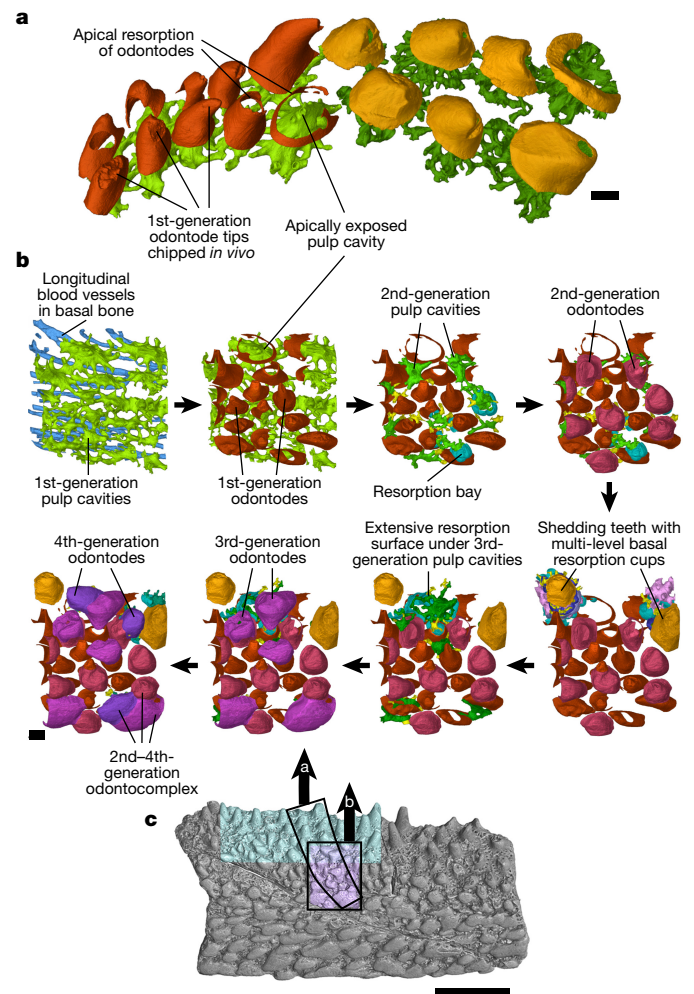
<sup>1</sup>Uppsala University, Department of Organismal Biology, Norbyvägen 18A, SE-752 36 Uppsala, Sweden. <sup>2</sup>Science for Life Laboratory and Uppsala University, Department of Organismal Biology, Norbyvägen 18A, SE-752 36 Uppsala, Sweden. <sup>3</sup>European Synchrotron Radiation Facility, 6 rue Jules Horowitz, 38043 Grenoble CEDEX, France.



**Figure 1 | The lower jaw of *Andreolepis*.** **a**, A block model from high-resolution PPC-SRμCT scan of LOG87-301DF, showing the internal architecture of the oral lamina. Oblique dorsolateral view. Scale bar, 0.1 mm. **b**, An overview of LOG87-301DF in lateral view, modelled from a low-resolution PPC-SRμCT scan. Dashed line indicates the transition between the facial lamina (which is parallel to the picture plane) and the oral lamina (which dips away from the viewer into the picture plane). Turquoise and purple highlighting indicates areas scanned at high resolution, black box with arrow indicates location of block shown in **a**. Scale bar, 1 mm. **c**, Simplified gnathostome phylogeny showing the probable position of *Andreolepis* and *Lophosteus*, based on refs 9 (overall topology, placement of *Lophosteus* in the osteichthyan stem group) and 8 (placement of *Andreolepis* in the osteichthyan stem group). *Andreolepis* is placed above *Lophosteus* because both it and crown Osteichthyes possess scale enamel<sup>1,22</sup>.

numerous buried resorption–overgrowth surfaces in the specimen that were not detected by the previous study. Second, that we have generated three-dimensional virtual models of the histology (Figs 1–3 and Extended Data Figs 1–5), rather than relying on virtual thin sections; this has uncovered a complex three-dimensional architecture of previously unseen embedded structures, allowing us to reconstruct the developmental events and functional history of the odontode skeleton in detail (Extended Data Fig. 6). The discovery of a sharply delineated dentition of true shedding teeth, each sitting on top of multiple stacked resorption cups representing episodes of shedding and replacement, strongly supports the original identification<sup>17</sup> of LOG87-301DF as a marginal jaw bone.

In cross-section, the jaw element has an inverted L-shape, with a horizontal oral lamina joined to a vertical facial lamina (Figs 1–3 and Extended Data Figs 1, 6). However, the boundary between the laminae is a transversely curved surface rather than a sharp edge; the odontode skeleton that we assign to the facial lamina (as we will discuss) in fact extends a short distance across this curve onto the horizontal surface. The base plate of dermal bone is covered externally by a complex layer of odontodes with pulp cavities, bone of attachment, vascular canals and buried resorption surfaces. The original extent of the base plate



**Figure 2 | The external odontode skeleton.** **a**, Two files of first-generation facial odontodes (red) and adjacent files of shedding teeth (gold) in oblique dorsolateral view. The rightmost shedding teeth occupy the lingual margin of the oral face of the bone; the labial margin of the oral face is marked approximately by the rightmost facial odontodes. The shedding teeth are exposed on the surface, but the first-generation facial odontodes are largely overgrown and have been dissected out digitally. **b**, Virtual dissection of the dorsal part of the facial lamina. A lateral view is shown, starting from the deepest layer (first-generation pulp cavities) and moving to progressively more superficial layers. Ventral is at the bottom of each image. The apically exposed pulp cavity in **a** and **b** is the same odontode. **c**, overview of LOG87-301DF showing location of detailed images **a** and **b**. Note that the tooth files in **a** are not strictly transverse, but oblique. Scale bars, 0.1 mm (**a** and **b**) and 1 mm (**c**). Colours as in Fig. 1a.

corresponds to a region of basal bone that lacks growth increments and contains a longitudinal vascular mesh (Figs 1a, 2b and Extended Data Fig. 1, vessels shown in light blue). The lingual margin of the tooth-bearing lamina has undergone later appositional growth, as shown by a series of growth increments in the basal bone (Fig. 1a, dark blue and Extended Data Fig. 2a, labelled) and a separate vascular supply for this region (Extended Data Fig. 1, vessels are shown pink and green).

The first generation of odontodes on the facial lamina, which lack underlying resorption surfaces, are low-crowned and longitudinally elongated on the ventral part of the lamina. On the dorsal curved surface, however, they become tall, conical and recurved (Fig. 2a, b, Extended Data Fig. 6). The transition between the two morphologies is gradual rather than abrupt (Fig. 2a, b). Many of the conical odontodes that are later completely overgrown by hard tissue preserve chipped and broken tips, readily distinguishable from resorption surfaces in



cross-section (Fig. 2a and Extended Data Fig. 3). This *in vivo* damage indicates that these odontodes had a biting function. They are organized in semi-regular, obliquely transverse files, with the overlapping relationship of their bases showing that new odontodes were added sequentially at the lingual margin of the odontode field (Fig. 2a, c and Extended Data Fig. 3). This odontode set extends a little way onto the horizontal lamina of the bone but stops short of its lingual edge, exposing a substantial rim (Extended Data Fig. 6a).

The later ontogenetic history of the jaw divides into two discrete spatial domains, one encompassing the facial lamina and the other the oral lamina (Extended Data Fig. 6). On the facial lamina, the first-generation odontodes become overgrown by several more generations that originate in the gaps between existing odontodes (Fig. 2b and Extended Data Fig. 6b–d). They form odontocomplexes ventrally, smoothing but hardly thickening the dentinal layer (Fig. 2b and Extended Data Fig. 4). On the dorsal curved surface, where the first-generation odontodes are tall and conical, subsequent generations remain taller than their ventral contemporaries but change through successive generations to an irregular, flat-topped morphology typical of dermal ornament. Near the lingual edge of the domain, these odontodes also have chipped tips, suggesting that they were not entirely covered with epithelium. Some overgrowing odontodes are associated with small, irregular bays of resorption (Fig. 2b and Extended Data Fig. 3), but these never develop into complete resorption cups.

The oral lamina develops differently (Figs 1a, 3 and Extended Data Fig. 6). After the first-generation odontodes have been at least partially overgrown by bone, there is a phase of bone resorption across the whole dorsal face of the horizontal lamina (Fig. 1a and Extended Data Fig. 6c, resorption surfaces shown in turquoise). Next, new bone and shedding teeth are added to this face (Extended Data Fig. 6d, e). The bone contains vasculature (Fig. 1a, shown in pink) that extends upwards from the longitudinal vascular mesh (pale blue) to the pulp cavities (green) of the shedding teeth (gold). These teeth all sit on large resorption cups, showing that osteoclast activity created a socket in the bone before the attachment of the tooth<sup>12,24</sup>. These resorption cups are not always perfectly centred under the functional teeth; they reflect the positions of the preceding shed teeth and show that the positioning of replacement teeth was flexible (Figs 1a, 3 and Extended Data Fig. 2a). The presence of up to three stacked resorption cups provides evidence of repeated basal resorption, shedding and replacement (Fig. 3d–f). Empty sockets are present but rare (Fig. 3a, b and Extended Data Fig. 5). The teeth form a semi-regular pattern (Figs 2a and 3a). The labial margin of the tooth field is irregular and in places partly overgrown by ornament, its position not necessarily corresponding with the lingual margin of the first-generation odontodes (Fig. 3b). Tooth replacement appears to be site-autonomous, not organized in a spatial pattern. The resorption cups of lingually positioned teeth tend to cut into those of labially positioned teeth, suggesting that teeth were generally added to the lingual margin of the field, although some exceptions have been observed (Fig. 3b). The shedding teeth on the labial margin of the oral lamina only have one or two resorption cups; tooth replacement seems to have been terminated by the partial overgrowth of non-shedding facial odontodes (Fig. 3a). The teeth on the lingual margin of the bone are larger than the rest and have, at most, two stacked resorption cups.

The shedding teeth differ histologically from the first-generation facial odontodes and the ornament odontodes that overgrow them. The dentine of the teeth is partly cellular and, when modelled in positive, their pulp cavities appear ‘warty’ owing to numerous projecting odontoblast lacunae. By contrast, the first-generation and ornament odontodes have dense, non-cellular dentine and smooth-walled pulp cavities (Extended Data Fig. 2a, c). The radial vessels of the shedding teeth, which pierce the bone of attachment and open onto the surface (Figs 1a, 3d–f and Extended Data Fig. 5b, c, shown in yellow), are also more numerous and regular in arrangement than those of other odontodes.

## Dental organization and ontogeny

The odontode organization of LOG87-301DF allows certain conclusions to be drawn about its ontogeny. The oral lamina of the jaw carries a dentition of true teeth, identifiable because multiple cycles of tooth shedding by basal resorption followed by replacement can be seen, with the replacement tooth on each occasion cemented into the resorption cup left by its predecessor (Fig. 3d–f). While this type of replacement is not universal for teeth<sup>4,12</sup>, it is unique to osteichthyan teeth and does not occur in other odontode categories.

The presence of stacked resorption cups indicates that replacement teeth were generated extraosseously, above the jaw bone. The fact that replacement was occurring in all parts of the dentition field, even in locations separated from the lingual margin of the jaw bone by up to three tooth rows (Fig. 3a, b), indicates that it did not involve a single lingually positioned dental lamina as seen in sharks or some tetrapods; it would have been impossible for replacement teeth generated at the lingual margin of the bone to move across such a field of fixed teeth to their final destination. Replacement teeth must have been generated adjacent to each standing tooth, as in the majority of actinopterygians, either directly from the epithelial folds of the standing teeth or from small successional dental laminae associated with them<sup>4–6,8–11</sup> (Extended Data Fig. 7).

The teeth are clearly demarcated from the facial odontodes. First, facial odontodes never sit on complete resorption cups, although small areas of resorption are associated with some of them; second, new facial odontodes are added in a gap-filling pattern; third, they are irregular and cushion-shaped; and fourth, younger odontodes overgrow older ones, suggesting that the facial odontodes never erupted through the epithelium. However, the first generation of facial odontodes does not fully conform to this description. They are organized in transverse files, are conical in shape and show *in vivo* tip-chipping that we interpret as biting damage. Similar damage is seen on the most lingual of the overgrowing facial odontodes, suggesting a persisting food-processing function.

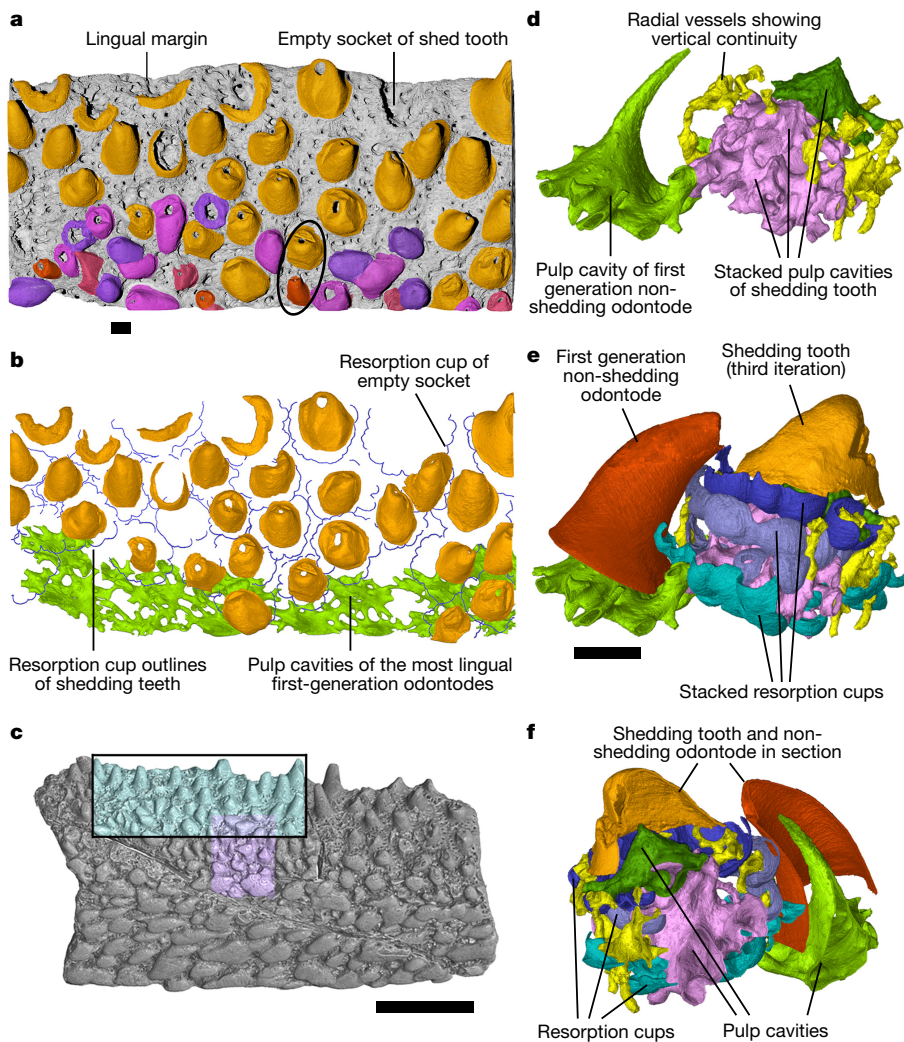
Our results also allow us to reinterpret data presented previously<sup>23</sup>, which the authors use to argue against the presence of true teeth in *Andreolepis*. Two factors seem to have caused them to interpret their scan incorrectly. First, as can be seen in figure 1b of ref. 23, NRM-PZ.P15910 has a much narrower field of shedding teeth than LOG87-301DF, comprising only one or two rows. As a result, most of the odontodes visible in the high-resolution figure 2d of ref. 23 belong to the external surface of the bone. Two overgrown first-generation facial odontodes are clearly visible in the middle of the image and only a single shedding tooth is present (on the far right). Second, while their scan shows overgrowth relationships fairly clearly, it does not capture resorption surfaces. The resorption cup underlying the shedding tooth in figure 2d of ref. 23 is almost invisible except at the right-hand margin of the bone and was not observed by the authors<sup>23</sup>. We therefore conclude that the dental organization of NRM-PZ.P15910 and LOG87-301DF is essentially identical.

## Origin of the osteichthyan dentition

The proposed phylogenetic position of *Andreolepis* as a stem osteichthyan<sup>1,16–19,22</sup> (Fig. 1c) is supported by its dentition. Tooth-shedding by basal resorption is a unique osteichthyan trait, absent in other gnathostomes, and supports assignment of *Andreolepis* into the osteichthyan total group. The broad tooth field and the lack of enamel, both on the teeth and on the facial odontodes, suggest that it belongs to the osteichthyan stem group. All early crown-group osteichthyans with marginal teeth have a well-defined tooth row, whereas an *Andreolepis*-like dentition is seen in the probable stem osteichthyan *Lophosteus*<sup>17</sup>, which is interpreted as less crownward than *Andreolepis* because it completely lacks enamel<sup>1,22</sup>. The dental organization of *Andreolepis* may therefore be primitive relative to that of crown osteichthyans.

The phylogenetic framework for interpreting the early evolution of vertebrate dentitions has recently become clearer thanks to several





**Figure 3 | The shedding dentition.** **a**, Dorsal view of the oral lamina with shedding dentition (gold), dermal ornament (purple, pink) and first-generation facial odontodes (red). Lingual margin at the top. **b**, Virtual dissection of the view in **a**, showing shedding teeth with outlines of resorption cups and pulp cavities of most lingual first-generation facial odontodes. Note that in **a**, very few of these first-generation odontodes are visible on the surface. **c**, An overview of LOG87-301DF, showing the location of **a** and **b**. **d–f**, Close-up images of a first-generation facial odontode and an adjacent shedding tooth in side-view. Structures are the same as those indicated with a black oval in **a**. Pulp cavities and vessels are shown in **d**; crowns and resorption cups in **e**; crowns and resorption cups cut through to reveal pulp cavities in **f**. Scale bars, 0.1 mm (**a**, **b** and **d–f**) and 1 mm (**c**). Colours as in Fig. 1a.

factors: the recognition that placoderms are a paraphyletic array of stem gnathostomes<sup>25–28</sup>; the discovery of *Entelognathus*, a maxillate placoderm showing that osteichthyan-like marginal jaw bones are primitive for crown gnathostomes<sup>27</sup>; and the reinterpretation of all acanthodians as stem chondrichthyans<sup>27–30</sup>. These form the basis for the scenario presented here (Extended Data Fig. 7).

Tooth addition and tooth replacement are functionally distinct phenomena that are sometimes conflated. Tooth addition involves the addition of new teeth as space is provided by the growth of the underlying jaw (examples include lungfish tooth plates, arthrodire gnathal plates and the dermal jaw bones of ischnacanthid acanthodians)<sup>31–36</sup>. In tooth replacement, new teeth are added at a faster rate than the jaw grows, and space is created for them by moving old teeth out of the way. The tubercular ‘dentition’ of *Entelognathus* shows no signs of shedding and replacement<sup>27</sup>. Neither is a shedding–replacement cycle known to occur on the gnathal plates of conventional placoderms<sup>32,33,35</sup>. Thus, jaw bones with a non-shedding, non-replacing dentition were presumably present at the gnathostome crown-group node.

Osteichthyan tooth replacement evolved through the introduction of a basal resorption–shedding process into this system; if *Andreolepis* represents the primitive osteichthyan condition, it follows that replacement teeth were originally produced extraosseously in association with each tooth position (Extended Data Fig. 7). The acanthodian–chondrichthyan type of tooth replacement, by contrast, evolved through substitution of the static dermal jaw bones with transversely arranged tooth whorls<sup>14,37–40</sup> (Extended Data Fig. 7). The occurrence of whorl-like sets of oropharyngeal denticles in some fossil jawless vertebrates<sup>14,41</sup> suggests that this involved co-option of an existing

patterning mechanism. In the phylogenetically deepest acanthodians, whorls are restricted to the jaw symphysis<sup>36</sup>; the system later extended to the entire jaws<sup>26,39</sup>. Parasymphysial tooth whorls also occur in early sarcopterygians<sup>25,42,43</sup> and may reflect similar co-option or possibly homology. In whorl-type tooth replacement, all new teeth are produced at the lingual margin of the dentition, suggesting the presence of a permanent dental lamina, either continuous or discontinuous<sup>44</sup>. Among tetrapods, a permanent dental lamina is present in both amphibians and amniotes<sup>4–6</sup> and may have evolved shortly after the water-to-land transition, in association with the loss of the coronoid dentition<sup>45</sup>.

The importance of the first-generation facial odontodes of *Andreolepis* remains to be considered. That the most linguallly positioned examples functioned as biting ‘teeth’ is evident. Yet labially–ventrally they grade into lower, anteroposteriorly elongated odontodes that extend some distance down the facial lamina, possibly all the way to the ventral margin. The lack of a sharp boundary suggests that even the most lingual first-generation odontodes are modified members of the dermal odontode set. Their rigid spatial organization, which contrasts with the gap-filling odontodes that overgrow them, can be matched in the scale odontodes of *Andreolepis*<sup>19</sup>.

The first-generation biting odontodes of *Andreolepis* have no precise equivalent in any known crown osteichthyan. Small, conical, apparently non-shedding odontodes are present labial to the shedding tooth row in *Psarolepis*<sup>1</sup> and many primitive actinopterygians<sup>46</sup>, but overgrowth relationships show that these are added late in ontogeny<sup>1</sup>. An ambiguous, partially tooth-like character is also seen in the positionally equivalent lip scales of ischnacanthid acanthodians<sup>36</sup>, basal members of the chondrichthyan stem group<sup>27–30</sup>. Together, these examples emphasize the

importance of early fossil gnathostomes in our understanding of dental evolution, not only as windows into the early history of the modern dentition but also as representatives of odontode systems outside the range of modern forms.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 16 April 2015; accepted 25 August 2016.**

**Published online 17 October 2016.**

- Qu, Q., Haitina, T., Zhu, M. & Ahlberg, P. E. New genomic and fossil data illuminate the origin of enamel. *Nature* **526**, 108–111 (2015).
- Fraser, G. J., Cerny, R., Soukup, V., Bronner-Fraser, M. & Streelman, J. T. The odontode explosion: the origin of tooth-like structures in vertebrates. *BioEssays* **32**, 808–817 (2010).
- Debiais-Thibaud, M. *et al.* The homology of odontodes in gnathostomes: insights from *Dlx* gene expression in the dogfish, *Scyliorhinus canicula*. *BMC Evol. Biol.* **11**, 307 (2011).
- Fraser, G. J., Graham, A. & Smith, M. M. Developmental and evolutionary origins of the vertebrate dentition: molecular controls for spatio-temporal organisation of tooth sites in osteichthyans. *J. Exp. Zool. B. Mol. Dev. Evol.* **306**, 183–203 (2006).
- Smith, M. M., Fraser, G. J. & Mitsiadis, T. A. Dental lamina as source of odontogenic stem cells: evolutionary origins and developmental control of tooth generation in gnathostomes. *J. Exp. Zool. B. Mol. Dev. Evol.* **312**, 260–280 (2009).
- Smith, M. M. & Johanson, Z. in *Great Transformations in Vertebrate Evolution* (eds Dial, K. P. *et al.*) 9–29. (Univ. Chicago Press, 2015).
- Grady, J. E. Tooth development in *Latimeria chalumnae* (Smith). *J. Morphol.* **132**, 377–387 (1970).
- Huysseune, A. Formation of a successional dental lamina in the zebrafish (*Danio rerio*): support for a local control of replacement tooth initiation. *Int. J. Dev. Biol.* **50**, 637–643 (2006).
- Huysseune, A. & Witten, P. E. An evolutionary view on tooth development and replacement in wild Atlantic salmon (*Salmo salar* L.). *Evol. Dev.* **10**, 6–14 (2008).
- Vandenplas, S., DeClercq, A. & Huysseune, A. Tooth replacement without a dental lamina: The search for epithelial stem cells in *Polypterus senegalus*. *J. Exp. Zool. B. Mol. Dev. Evol.* **322**, 281–293 (2014).
- Smith, M. M. *et al.* Making teeth to order: conserved genes reveal an ancient molecular pattern in paddlefish (Actinopterygii). *Proc. Biol. Sci.* **282**, 20142700 (2015).
- Witten, P. E. & Huysseune, A. A comparative view on mechanisms and functions of skeletal remodelling in teleost fish, with special emphasis on osteoclasts and their function. *Biol. Rev. Camb. Philos. Soc.* **84**, 315–346 (2009).
- Reif, W.-E. in *Evolutionary Biology* (eds Hecht, M. K. *et al.*) 287–368 (Plenum Press, 1982).
- Smith, M. M. & Coates, M. I. in *Major Events in Early Vertebrate Evolution: Palaeontology, Phylogeny, Genetics and Development* (ed. Ahlberg, P. E.) 223–240 (Taylor & Francis, 2001).
- Gross, W. Fragliche Actinopterygier-Schuppen aus dem Silur Gotlands. *Lethaia* **1**, 184–218 (1968).
- Janvier, P. On the oldest known teleostome fish *Andreolepis hedei* Gross (Ludlow of Gotland), and the systematic position of the lophosteids. *ENSV TA Toim. Geol.* **27**, 86–95 (1978).
- Botella, H., Blom, H., Dorka, M., Ahlberg, P. E. & Janvier, P. Jaws and teeth of the earliest bony fishes. *Nature* **448**, 583–586 (2007).
- Chen, D., Janvier, P., Ahlberg, P. E. & Blom, H. Scale morphology and squamation of the Late Silurian osteichthyan *Andreolepis* from Gotland, Sweden. *Hist. Biol.* **24**, 411–423 (2012).
- Qu, Q., Sanchez, S., Blom, H., Tafforeau, P. & Ahlberg, P. E. Scales and tooth whorls of ancient fishes challenge distinction between external and oral 'teeth'. *PLoS One* **8**, e71890 (2013).
- Fredholm, D. Vertebrates in the Ludlovian Hemse Beds of Gotland, Sweden. *Geol. Foren. Stockholm Forh.* **110**, 157–179 (1988).
- Sire, J.-Y., Donoghue, P. C. J. & Vickaryous, M. K. Origin and evolution of the integumentary skeleton in non-tetrapod vertebrates. *J. Anat.* **214**, 409–440 (2009).
- Friedman, M. & Brazeau, M. D. A reappraisal of the origin and basal radiation of the Osteichthyes. *J. Vertebr. Paleontol.* **30**, 36–56 (2010).
- Cunningham, J. A., Rücklin, M., Blom, H., Botella, H. & Donoghue, P. C. J. Testing models of dental development in the earliest bony vertebrates, *Andreolepis* and *Lophosteus*. *Biol. Lett.* **8**, 833–837 (2012).
- Wistuba, J., Bolte, M. & Clemen, G. The odontoclasts of *Ambystoma mexicanum*. *Ann. Anat.* **182**, 415–422 (2000).
- Zhu, M., Wang, W. & Yu, X. in *Morphology, Phylogeny and Paleobiogeography of Fossil Fishes* (eds Elliott, D. K. *et al.*) 199–214 (Verlag Dr. Friedrich Pfeil, 2010).
- Brazeau, M. D. The braincase and jaws of a Devonian 'acanthodian' and modern gnathostome origins. *Nature* **457**, 305–308 (2009).
- Zhu, M. *et al.* A Silurian placoderm with osteichthyan-like marginal jaw bones. *Nature* **502**, 188–193 (2013).
- Dupret, V., Sanchez, S., Goujet, D., Tafforeau, P. & Ahlberg, P. E. A primitive placoderm sheds light on the origin of the jawed vertebrate face. *Nature* **507**, 500–503 (2014).
- Giles, S., Friedman, M. & Brazeau, M. D. Osteichthyan-like cranial conditions in an Early Devonian stem gnathostome. *Nature* **520**, 82–85 (2015).
- Brazeau, M. D. & de Winter, V. The hyoid arch and braincase anatomy of *Acanthodes* support chondrichthyan affinity of 'acanthodians'. *Proc. R. Soc. B* **282**, 20152210 (2015).
- Smith, M. M. & Krupina, N. I. Conserved developmental processes constrain evolution of lungfish dentitions. *J. Anat.* **199**, 161–168 (2001).
- Rücklin, M. *et al.* Development of teeth and jaws in the earliest jawed vertebrates. *Nature* **491**, 748–751 (2012).
- Young, G. C., Lelièvre, H. & Goujet, D. Primitive jaw structure in an articulated brachyothracid arthrodire (placoderm fish; Early Devonian) from Southeastern Australia. *J. Vertebr. Paleontol.* **21**, 670–678 (2001).
- Ørvig, T. Acanthodian dentition and its bearing on the relationships of the group. *Palaeontogr. Abt. A* **143**, 119–150 (1973).
- Gross, W. Über das Gebiss der Acanthodier und Placodermen. *Zool. J. Linn. Soc. London* **47**, 121–130 (1967).
- Blais, S. A., MacKenzie, L. A. & Wilson, M. V. H. Tooth-like scales in Early Devonian eugnathostomes and the 'outside-in' hypothesis for the origins of teeth in vertebrates. *J. Vertebr. Paleontol.* **31**, 1189–1199 (2011).
- Williams, M. Tooth retention in cladodont sharks: with a comparison between primitive grasping and swallowing, and modern cutting and gouging feeding mechanisms. *J. Vertebr. Paleontol.* **21**, 214–226 (2001).
- Miles, R. Articulated acanthodian fishes from the Old Red Sandstone of England, with a review of the structure and evolution of the acanthodian shoulder-girdle. *Bull. Br. Mus. Nat. Hist. (Geol.)* **24**, 111–213 (1973).
- Brazeau, M. D. A revision of the anatomy of the Early Devonian jawed vertebrate *Ptomacanthus anglicus* Miles. *Palaeontology* **55**, 355–367 (2012).
- Hanke, G. F. & Wilson, M. V. H. Anatomy of the Early Devonian acanthodian *Brochoadmones milesi* based on nearly complete body fossils, with comments on the evolution and development of paired fins. *J. Vertebr. Paleontol.* **26**, 526–537 (2006).
- Rücklin, M., Giles, S., Janvier, P. & Donoghue, P. C. Teeth before jaws? Comparative analysis of the structure and development of the external and internal scales in the extinct jawless vertebrate *Loganellia scotica*. *Evol. Dev.* **13**, 523–532 (2011).
- Jarvik, E. Middle and Upper Devonian Porolepiformes from East Greenland with special reference to *Glyptolepis groenlandica* n. sp. *Meddr. Grønland* **187**, 1–295 (1972).
- Andrews, S. M., Long, J., Ahlberg, P., Barwick, R. & Campbell, K. The structure of the sarcopterygian *Onychodus jandemarrai* n. sp. from Gogo, Western Australia: with a functional interpretation of the skeleton. *Trans. R. Soc. Edinb. Earth Sci.* **96**, 197–307 (2005).
- Smith, M. M. & Johanson, Z. in *From Clone to Bone: The Synergy of Morphological and Molecular Tools in Paleobiology* (eds Asher, R. J. & Müller, J.) 166–206 (Cambridge Univ. Press, 2012).
- Ahlberg, P. E. & Clack, J. A. Lower jaws, lower tetrapods - a review based on the Devonian genus *Acanthostega*. *Trans. R. Soc. Edinb. Earth Sci.* **89**, 11–46 (1998).
- Gardiner, B. G. The relationships of the palaeoniscid fishes, a review based on new specimens of *Mimia* and *Moythomasia* from the Upper Devonian of Western Australia. *Bull. Br. Mus. Nat. Hist.* **37**, 173–428 (1984).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank P. Janvier for originally bringing LOG87-301DF to our attention and Lund University for the loan of the specimen. Scanning was performed at the ESRF as part of proposal ES 151 and with in-house beamtime provided to S.S. during her time as a postdoctoral fellow at ESRF. D.C., P.E.A. and S.S. thank ERC (Advanced Investigator Grant 233111) for support. D.C. and P.E.A. also thank the Knut and Alice Wallenberg Foundation (Wallenberg Scholarship awarded to P.E.A.).

**Author Contributions** The project was conceived by D.C., P.E.A. and H.B. Scanning and reconstruction of datasets was carried out by S.S. and P.T., with minor contributions to the scanning by D.C. and P.E.A. All segmentation (morphological modelling) of datasets was done by D.C., with technical support by S.S. Morphological and developmental interpretation of the models was made by D.C. in collaboration with P.E.A. and H.B. The text was written by P.E.A. with contributions from D.C. and other authors. All figure images and animations were by D.C., except Extended Data Figs 6 and 7, which were by P.E.A. Figures assembled by P.E.A. and D.C. with input from other authors.

**Author Information** The synchrotron data will be made available through the European Synchrotron Radiation Facility (ESRF) palaeontology database (<http://paleo.esrf.eu>), catalogued by taxon name and publication reference. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.C. ([donglei.chen@ebc.uu.se](mailto:donglei.chen@ebc.uu.se)) or P.E.A. ([per.ahlberg@ebc.uu.se](mailto:per.ahlberg@ebc.uu.se)).

**Reviewer Information** Nature thanks M. Coates, M. M. Smith and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

**Acquisition and reconstruction of the PPC–SRmCT dataset.** A jaw bone of *Andreolepis hedei*, previously freed from the rock with dilute acetic acid, was imaged at beamline ID19 of the ESRF, using PPC–SRμCT. A multi-resolution approach was used. The whole fragment was scanned at medium resolution (voxel size: 5.06 μm), with two high-resolution scans (voxel size: 0.678 and 0.696 μm) targeted on regions of interest.

The medium-resolution scan was performed using a filtered polychromatic beam. The distance between the sample and the camera, mounted on the optical system, was 400 mm. The camera was a FreLoN 2K14 detector (fast-readout low-noise camera<sup>47</sup>), mounted on a lenses-based magnification system giving an isotropic voxel size of 5.06 μm. The scintillator was a 200-μm thick LuAG:Ce (lutetium aluminium garnet doped with cerium). The experiment was performed with a W150 wiggler (11 dipoles, 150 mm period) with a gap opened to 87 mm. The beam was filtered with 2 mm of aluminium. The resulting detected spectrum had an average energy of 30.2 keV (18–35 keV FWHM). The time of exposure was of 0.3 s. 1,999 projections were taken over 360°.

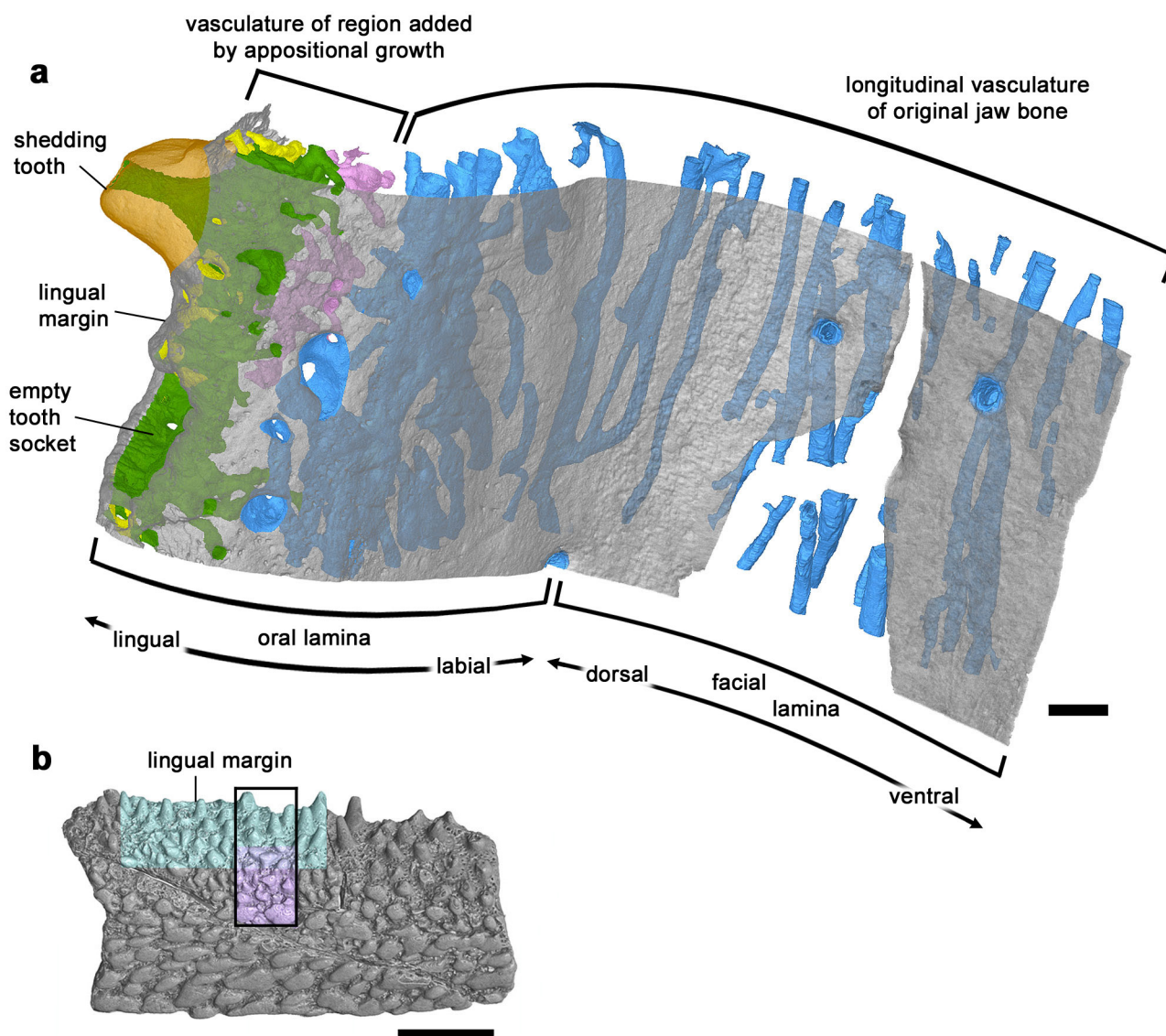
The high-resolution scans were done in two separate experiments. The first scan was done with a voxel size of 0.678 μm (obtained with an objective 10×, NA 0.3, coupled with a 2× eyepiece), with the sample set up at a distance of 30 mm from the optics. The experiment was performed with a monochromatic beam obtained using a single-bounce 2.5-nm period W/B4C multilayer monochromator with a fixed energy of 30 keV. The gap of the undulator of the insertion device (U32U) was set to 12.38 mm. The optics, associated to a europium-doped gadolinium gallium garnet crystal of 10 mm thickness (GGG:Eu10) scintillator, were coupled to a FreLoN 2K14 detector<sup>47</sup>. During the scan, 2,000 projections were taken over 180°, with a time of exposure of 0.3 s.

The second high-resolution scan was done with a voxel size of 0.696 μm (obtained with an objective 10×, NA 0.3, coupled with a 2× eyepiece, a GGG10 scintillator and a FreLoN 2k14 detector). The sample was set up at a distance of 15 mm from the optics. The gap of the undulator U13 was set to 20 mm and provided a pink beam (direct beam with a single main narrow harmonic) at an energy of 26.5 keV. The time of exposure was of 0.3 s. A total of 2,499 projections were taken over 360°.

All scanning data were reconstructed using a classical filtered back-projection algorithm (PyHST software, ESRF) adapted to local tomography<sup>48</sup>. This was done in edge-detection mode or with a modified version<sup>49</sup> of a single-distance phase retrieval approach<sup>50</sup>, based on an assumption of chemical homogeneity. Artefacts (for example, rings or movements) were corrected on reconstructed slices. The stacks of 2,048 TIFF images (Supplementary Videos 1, 2) were modelled in 3D using VG Studio MAX 2.2 (Volume Graphics) software.

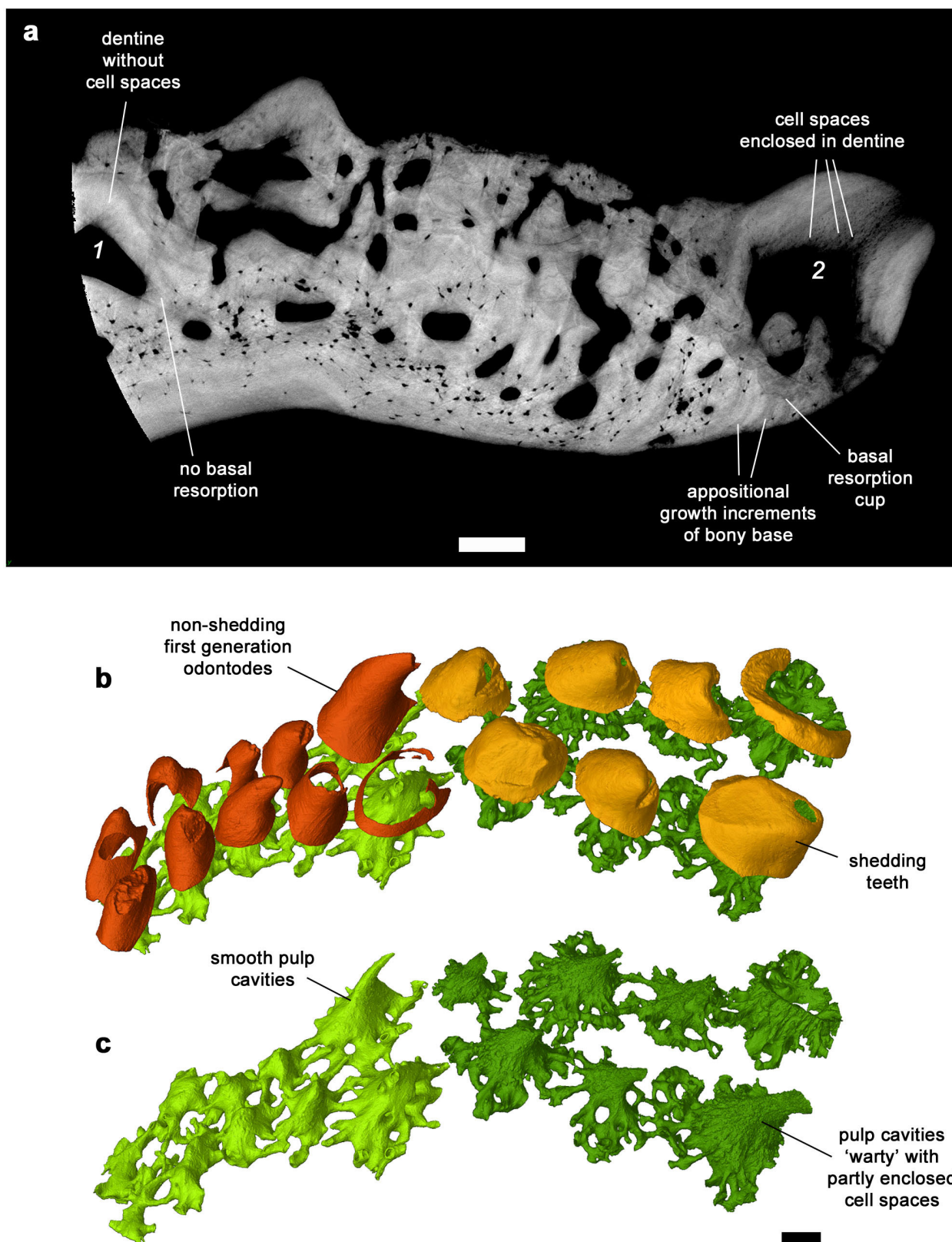
47. Labiche, J.-C. *et al.* Invited article: the fast readout low noise camera as a versatile x-ray detector for time resolved dispersive extended x-ray absorption fine structure and diffraction studies of dynamic problems in materials science, chemistry, and catalysis. *Rev. Sci. Instrum.* **78**, 091301–091311 (2007).
48. Lak, M. *et al.* Phase contrast X-ray synchrotron imaging: opening access to fossil inclusions in opaque amber. *Microsc. Microanal.* **14**, 251–259 (2008).
49. Sanchez, S., Ahlberg, P. E., Trinajstić, K. M., Mirone, A. & Tafforeau, P. Three-dimensional synchrotron virtual paleohistology: a new insight into the world of fossil bone microstructures. *Microsc. Microanal.* **18**, 1095–1105 (2012).
50. Paganin, D., Mayo, S. C., Gureyev, T. E., Miller, P. R. & Wilkins, S. W. Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object. *J. Microsc.* **206**, 33–40 (2002).





**Extended Data Figure 1 | Vasculature of the jaw bone. a,** A block model from an oblique posteromedial view showing the internal face of the jaw bone (rendered semi-transparent), the vasculature and a single shedding tooth on the lingual margin. Note the abrupt change in vascular

organization between the original bone and the part added by appositional growth. Colours as in Fig. 1a. **b,** overview of LOG87-301DF showing location of a. Scale bars, 0.1 mm (a) and 1 mm (b).

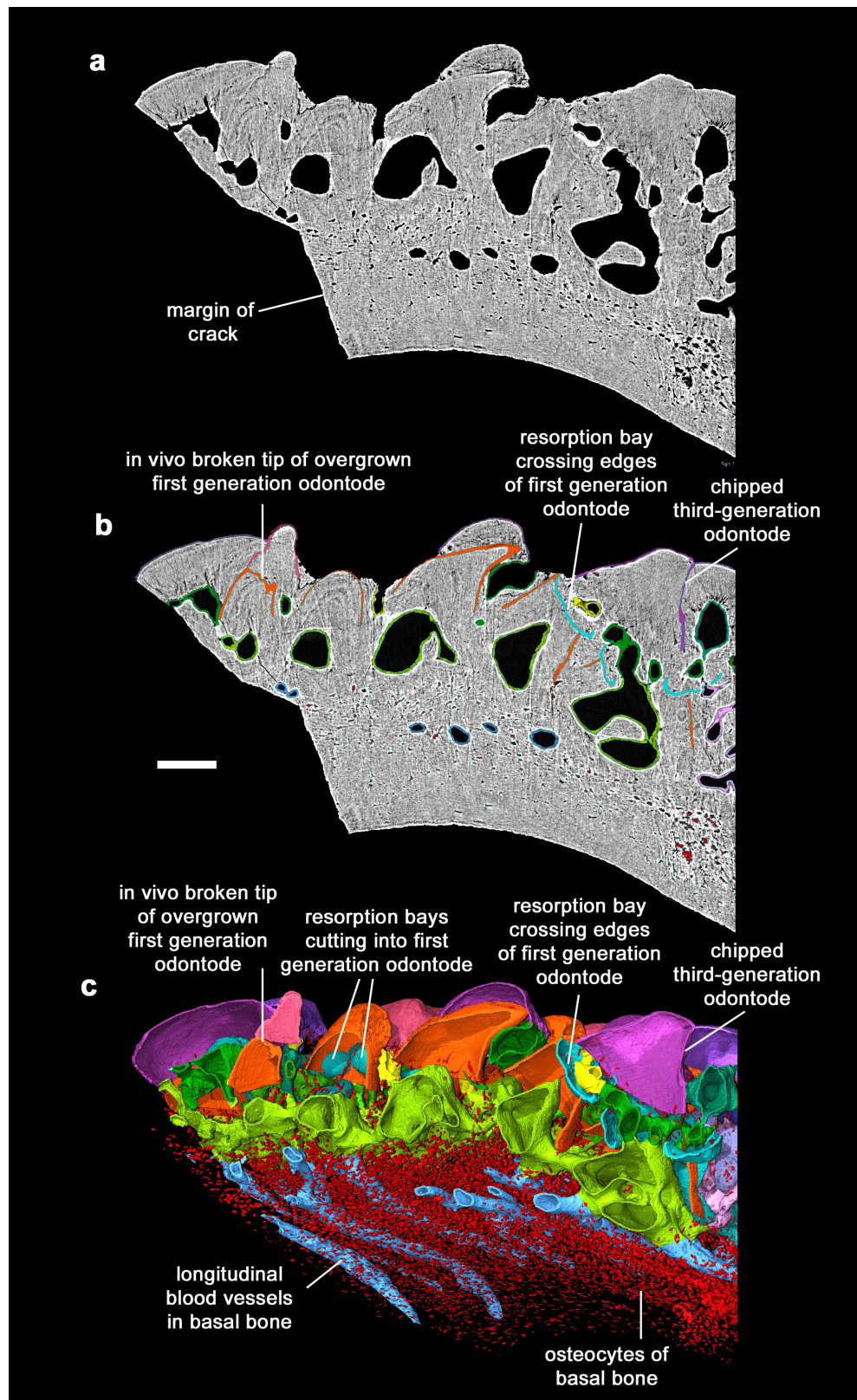


#### Extended Data Figure 2 | Odontodes and growth increments in section.

**a**, A virtual section from the second high-resolution dataset (turquoise in Figs 1b, 2c, 3c), showing a transverse section through tooth-bearing region. On the left, 1 indicates a buried first-generation facial odontode with a characteristic smooth-sided pulp cavity and purely tubular dentine. On the right, 2 indicates a shedding tooth with basal resorption cup and

partly cellular dentine. **b**, **c**, Two tooth files of first-generation odontodes and shedding teeth (see also Fig. 2a); **c** shows the morphology of the pulp cavities. Note the marked difference in surface texture: facial odontodes have smooth pulp cavities, but in shedding teeth the pulp cavities appear 'warty' because of numerous attached cell lacunae. Scale bars, 0.1 mm.



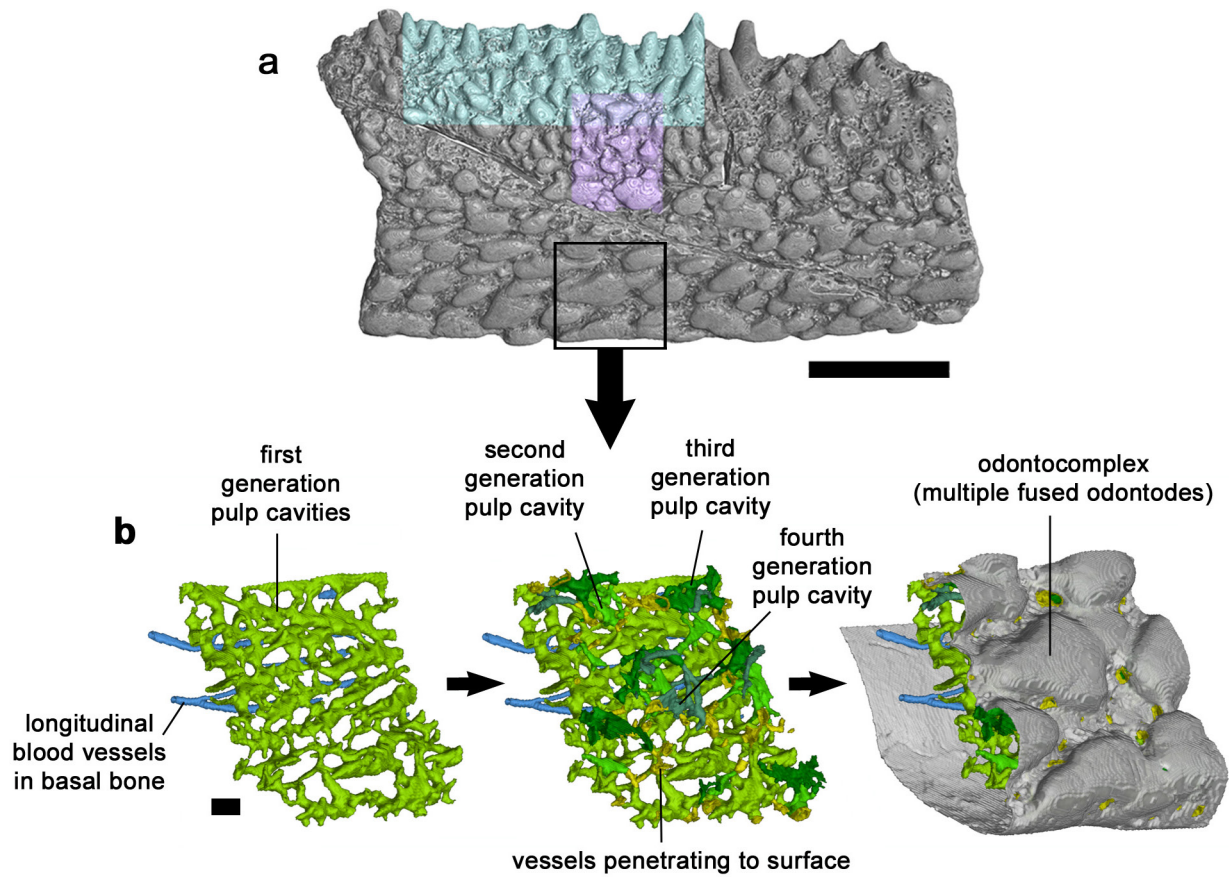


### Extended Data Figure 3 | Breakage and resorption of odontodes.

**a, b**, Virtual sections from the first high-resolution scan dataset (mauve in Figs 1b, 2c, 3c), illustrating examples of odontode damage by *in vivo* breakage (buried by later overgrowth) and resorption. **c**, A block model of the same region, with bone rendered transparent, terminating at a slice plane corresponding to virtual section in **a** and **b**. The ventral aspect is on the left in these images; the lateral surface of the bone is dorsal. Panel **a** is

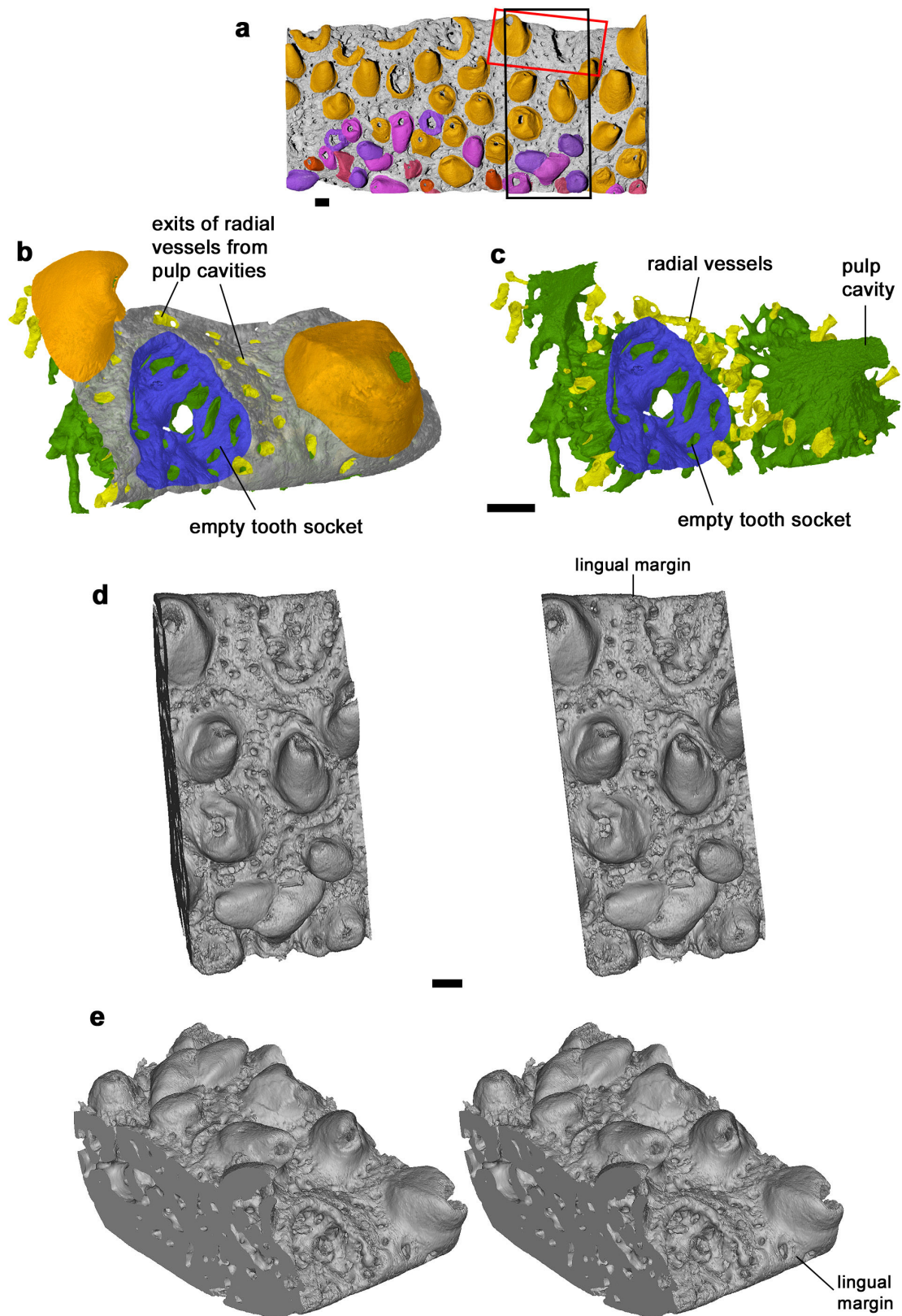
shown without labelling or highlighted model surfaces, to facilitate reader evaluation of the data quality. In **b**, modelled surfaces are shown, using the same colour code as Fig. 1a. Resorption can be distinguished from breakage because the characteristically curved resorption bays cut through both the odontode and the surrounding bony tissue that embeds it. Note that the facial odontodes all lack basal resorption cups. Scale bar, 0.1 mm.





**Extended Data Figure 4 | Odontodes on the ventral part of the facial lamina. a,** An overview of LOG87-301DF showing location of detail images in **b**. **b,** Series of histological models, moving from deeper on the left to more superficial on the right. Shown are the vascular supply of the

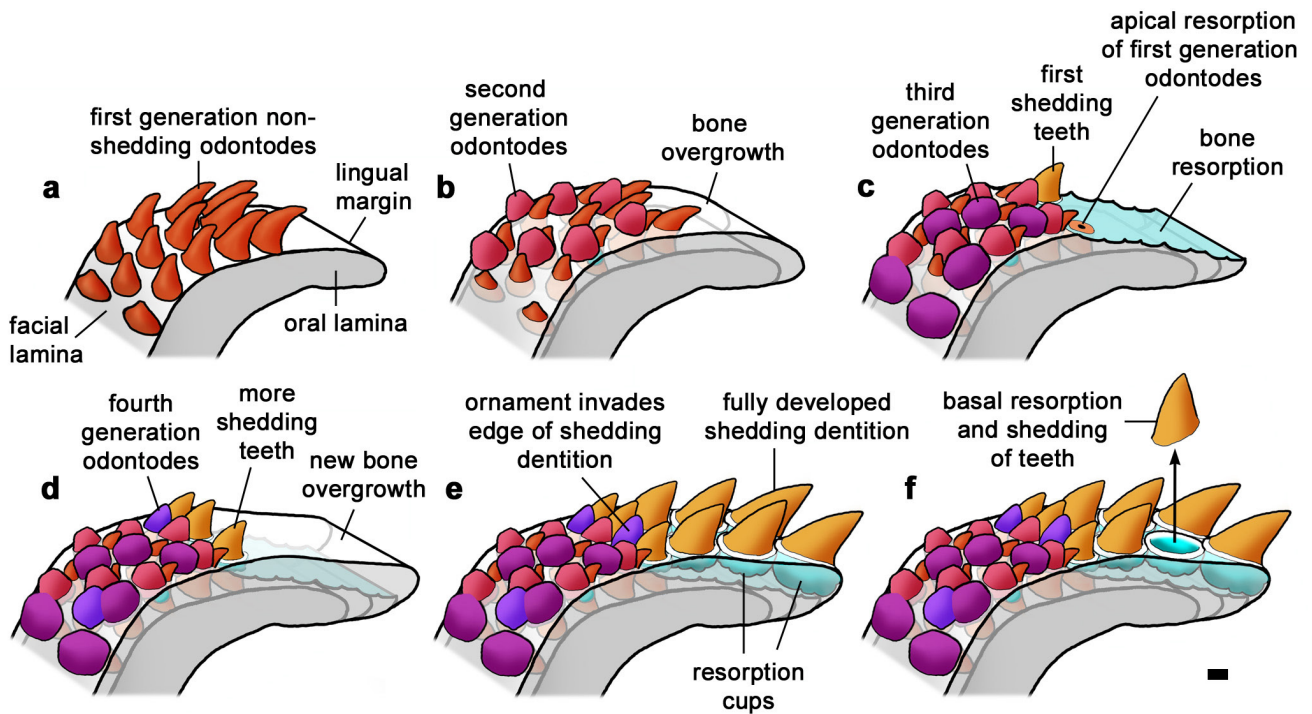
basal plate, four generations of odontode pulp cavities and the exposed surface of odontocomplexes. Colours of internal structures are as in Fig. 1a. Scale bars, 1mm (**a**), 0.1 mm (**b**).



#### Extended Data Figure 5 | Morphology of an empty tooth socket.

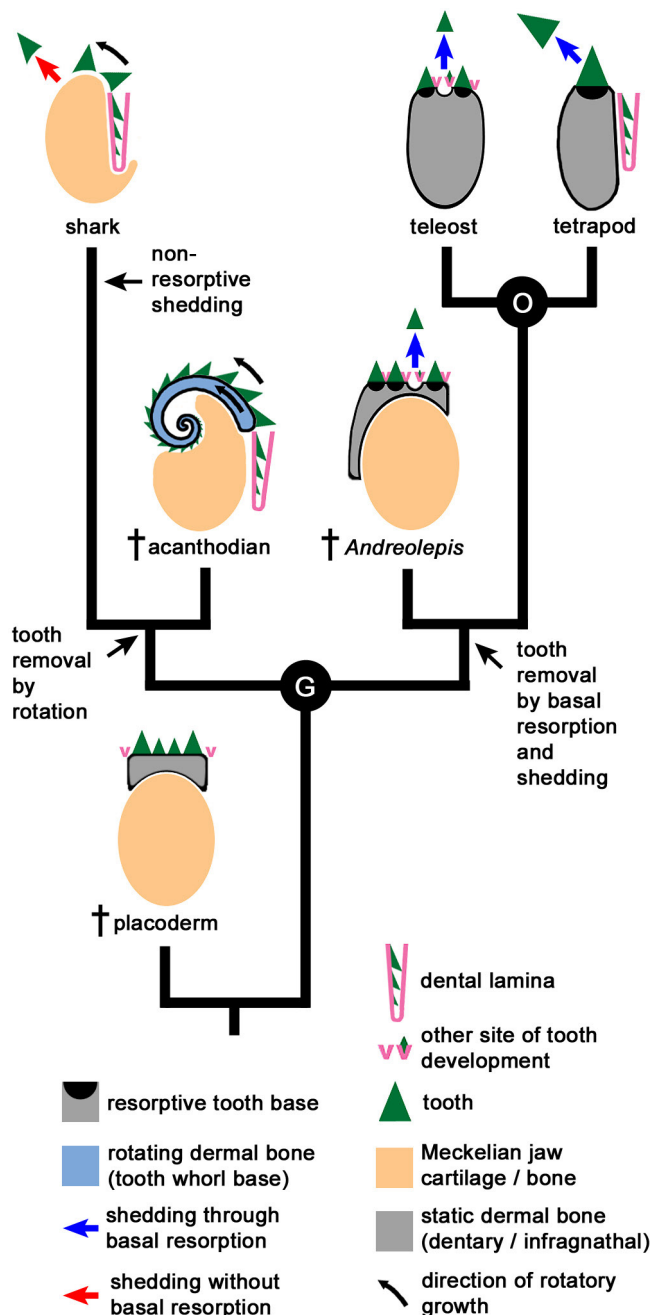
**a**, Area of shedding dentition in dorsal view (see also Fig. 3a). The red box indicates the area shown in **b** and **c**, the black box indicates area shown in **d** and **e**. **b**, Empty tooth socket (see also Fig. 3a) and two adjacent teeth in oblique anteromesial view, socket indicated in dark blue. **c**, The same region, rendered to show only pulp cavities and radial vessels.

**d, e**, Stereo pairs in dorsal (**d**) and oblique anteromesiodorsal (**e**) views. Note the distinct raised lip of the empty tooth socket. This socket contains no dentine remnants and is an exposed example of the structures elsewhere described as basal resorption cups (Fig. 3e, f and Extended Data Fig. 2a). Scale bars, 0.1 mm.



**Extended Data Figure 6 | Growth model.** a–f, Block models of the dorsal part of the bone in oblique dorsolateral view, lingual margin to the right, showing reconstructed ontogeny of the dentition and dermal ornament. Images are slightly schematized, vascular canals not shown. Scale bar, 0.1 mm. Colours as in Fig. 1a.





#### Extended Data Figure 7 | The evolution of tooth replacement.

Simplified gnathostome phylogeny, based on recent publications<sup>1,17,22,26–29</sup>, showing diagrammatic cross-sections of lower jaws. The labial aspect is to the left and the lingual to the right in each diagram. Crosses denote fossil taxa. Sites of tooth development (pink) are known in extant taxa and inferred for fossil taxa. A permanent dental lamina can be either continuous along the length of the jaw, or discontinuous and associated to an individual tooth position. Other sites of tooth development include transient dental laminae and direct development from the oral epithelium<sup>2–11</sup>. The placoderm is based on a primitive arthrodire<sup>33</sup>, the acanthodian on a tooth-whorl-bearing form such as *Ptomacanthus*<sup>38,39</sup>. The actinopterygian and tetrapod examples represent common modes of tooth replacement in the respective groups. Both clades also contain different styles of tooth replacement, including intraosseous dental laminae<sup>4–6,8–11</sup>. Note that acanthodian-like parasymphysial tooth whorls occur in some fossil osteichthyans<sup>25,42,43</sup>, if these are homologous with those of chondrichthyans (a matter of current debate), then tooth removal by rotation would map to node G. Node O denotes the osteichthyan crown group node, node G the gnathostome crown group node.

# Evolution of Osteocrin as an activity-regulated factor in the primate brain

Bulent Ataman<sup>1\*</sup>, Gabriella L. Boulting<sup>1\*</sup>, David A. Harmin<sup>1</sup>, Marty G. Yang<sup>1</sup>, Mollie Baker-Salisbury<sup>1</sup>, Ee-Lynn Yap<sup>1</sup>, Athar N. Malik<sup>1</sup>, Kevin Mei<sup>1</sup>, Alex A. Rubin<sup>1</sup>, Ivo Spiegel<sup>1</sup>, Ershele Duresse<sup>1</sup>, Nikhil Sharma<sup>1</sup>, Linda S. Hu<sup>1</sup>, Mihovil Pletikos<sup>2</sup>, Eric C. Griffith<sup>1</sup>, Jennifer N. Partlow<sup>3</sup>, Christine R. Stevens<sup>4</sup>, Mazhar Adli<sup>5</sup>, Maria Chahrour<sup>6</sup>, Nenad Sestan<sup>2</sup>, Christopher A. Walsh<sup>3</sup>, Vladimir K. Berezovskii<sup>1</sup>, Margaret S. Livingstone<sup>1</sup> & Michael E. Greenberg<sup>1</sup>

**Sensory stimuli drive the maturation and function of the mammalian nervous system in part through the activation of gene expression networks that regulate synapse development and plasticity. These networks have primarily been studied in mice, and it is not known whether there are species- or clade-specific activity-regulated genes that control features of brain development and function. Here we use transcriptional profiling of human fetal brain cultures to identify an activity-dependent secreted factor, Osteocrin (OSTN), that is induced by membrane depolarization of human but not mouse neurons. We find that OSTN has been repurposed in primates through the evolutionary acquisition of DNA regulatory elements that bind the activity-regulated transcription factor MEF2. In addition, we demonstrate that OSTN is expressed in primate neocortex and restricts activity-dependent dendritic growth in human neurons. These findings suggest that, in response to sensory input, OSTN regulates features of neuronal structure and function that are unique to primates.**

The complexity of primate cognition and behaviour is thought to be the culmination of clade- and species-specific aspects of brain ontogenesis. In anthropoid primates, especially humans, this process results in a hugely expanded cerebral cortex, an increase in morphological and cellular diversity, a vast increase in the complexity of neuronal connectivity, and a prolonged period of experience-driven circuit development<sup>1,2</sup>. An unresolved question is how distant mammals, such as mice and humans, can have nearly the same number of genes when the human brain is so vastly more complex in form and function than the mouse brain. Recent studies have begun to address this question by elucidating the underlying genetic and molecular mechanisms that contribute to primate brain development, focusing on features of neural stem and progenitor cell biology in primates and in mice<sup>3–5</sup>. However, there has been limited progress in identifying the specific genes that underlie the complexity of neuronal connectivity in primates<sup>6</sup> or the genes that regulate primate-specific aspects of experience-dependent brain development. In this study, we used RNA sequencing (RNA-seq) to characterize the activity-regulated transcriptome in human fetal brain neurons and thereby to determine whether features of this signalling network have evolved specifically in humans or primates in general and might control aspects of experience-dependent brain development. We report that the non-neuronal mouse gene *Ostn* has been repurposed via evolutionary changes that give rise to a new enhancer sequence, which in response to sensory experience drives *OSTN* expression in the primate brain and may restrict dendritic growth in the developing cortex.

## Activity-regulated neuronal transcriptome

We established a dissociated culture system for primary human fetal brain cultures (hFBCs) (Extended Data Fig. 1a). Although these cultures exhibit significant cellular heterogeneity (Extended Data Fig. 1b, c), cultures obtained from independent brain samples nevertheless

have reproducible gene expression profiles (Extended Data Fig. 1f) that cluster more closely with those of brain tissue than of other human tissues (Extended Data Fig. 1g). By gene expression analysis (Extended Data Fig. 2a–f) and immunostaining for cortical markers (Extended Data Fig. 1d, e), we confirmed that hFBCs are enriched for cortical neuronal subtypes.

We induced voltage-dependent calcium influx into hFBCs by exposing the cultures to elevated levels of potassium chloride (55 mM KCl) to trigger synchronous membrane depolarization, which mimics the cellular response to neuronal activity. This method has been shown to reliably induce a pattern of activity-dependent gene transcription in cultured rodent neurons<sup>7,8</sup> that is highly similar to the pattern of gene induction that occurs in the intact brain in response to a wide variety of physiological stimuli<sup>9</sup>. Activity-induced transcriptional responses in hFBCs at the two chosen time points are consistent with the well-established biphasic transcriptional program observed in rodent neurons<sup>8</sup>. Early-response genes (ERGs) encoding transcriptional regulators, including the well-characterized immediate-early genes *NPAS4*, *FOS*, *JUNB*, *NR4A1*, *NR4A3*, and *FOSB* (Extended Data Fig. 3a, c), were induced within 1 h of depolarization, and a larger set of late-response genes (LRGs) was then induced within 6 h of membrane depolarization (Extended Data Fig. 3b). The LRGs preferentially encode secreted, cytoplasmic, and transmembrane factors (Extended Data Fig. 3d), including previously reported activity-regulated loci<sup>8</sup> such as *BDNF*, *ADCYAP1*, *PCSK1* and *GPR22* (Supplementary Table 1).

Although we found that the activity-dependent transcriptomes of rodent and human cultures shared many common induced genes (Supplementary Tables 2–4), we identified a group of genes that were selectively induced in response to membrane depolarization in hFBCs. These included the secreted protein-encoding gene *OSTN*, and two primate-specific genes: *ZNF331*, a kruppel-associated box zinc-finger

<sup>1</sup>Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>2</sup>Department of Neuroscience and Kavli Institute for Neuroscience, Yale School of Medicine, New Haven, Connecticut 06510, USA. <sup>3</sup>Division of Genetics and Genomics, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>4</sup>Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. <sup>5</sup>University of Virginia, School of Medicine, Department of Biochemistry and Molecular Genetics, Charlottesville, Virginia 22903, USA. <sup>6</sup>McDermott Center for Human Growth and Development, Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

\*These authors contributed equally to this work.

protein, and the long non-coding RNA *LINC00473* (Extended Data Fig. 3a, b). The mouse *Ostn* gene (also known as *Musclin*) encodes a small (12-kD) secreted protein that is expressed selectively in mouse skeletal muscle and bone, but not in the mouse brain<sup>10–12</sup>. To investigate the functional significance of differences in activity-dependent gene expression between mice and humans, we focused our efforts on *OSTN*, as it was the most highly induced gene in our transcriptome analysis.

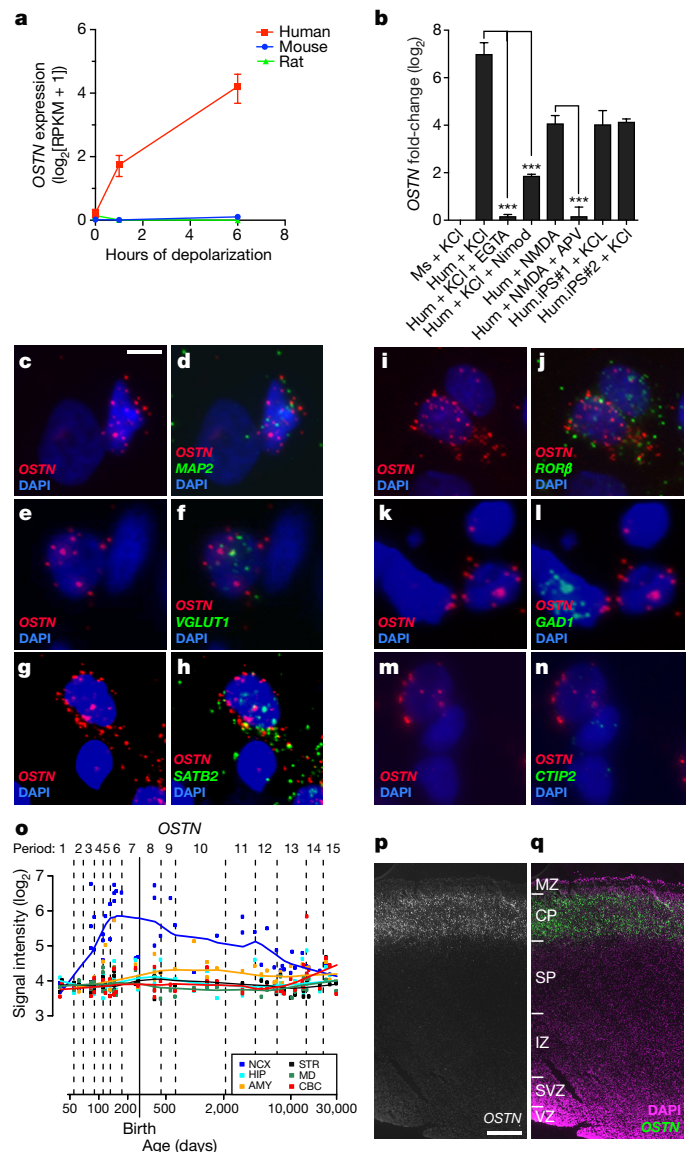
### *OSTN* is expressed in the human cortex

RNA-seq (Fig. 1a) and quantitative PCR with reverse transcription (RT-PCR) analysis confirmed that *OSTN* mRNA was significantly induced (>100-fold) in human fetal brain cultures in a manner that is dependent on calcium influx through L-type voltage-sensitive calcium channels (Fig. 1b); however, *Ostn* mRNA was not expressed or induced in mouse or rat neurons (Fig. 1a, b and Extended Data Fig. 4). *OSTN* mRNA was also induced upon exposure of hFBCs to the glutamate receptor agonist *N*-methyl-D-aspartate (NMDA) (Fig. 1b). The induction of *OSTN* mRNA is not a singular feature of hFBCs, inasmuch as *OSTN* mRNA was also induced upon membrane depolarization in two other human neuronal culture systems (Fig. 1b) derived from human induced pluripotent stem cells (iPSCs) (Extended Data Fig. 5). We next used fluorescence *in situ* hybridization (FISH) analysis to identify the neuronal subtypes that expressed *OSTN* mRNA. This analysis revealed that *OSTN* mRNA induction in hFBCs occurred selectively in a subset (*SATB2*<sup>+</sup>, *RORB*<sup>+</sup>, *BCL11B*<sup>+</sup> (also known as *CTIP2*)) of glutamatergic excitatory neurons (*MAP2*<sup>+</sup>, *SLC17A7*<sup>+</sup> (also known as *VGLUT1*), *GAD1*<sup>−</sup>) (Fig. 1c–n).

These culture-based findings are corroborated by data from the available Human Brain Transcriptome (HBT) database<sup>13</sup> and the BrainSpan Project (<http://www.brainspan.org/>). Analysis of these data revealed that *OSTN* mRNA is enriched across human neocortical regions, but appears not to be expressed at appreciable levels in other compartments of the neocortical wall or in brain regions such as the hippocampus, striatum, mediodorsal nucleus of the thalamus and cerebellum (Fig. 1o and Extended Data Fig. 6). *OSTN* expression increases during the course of fetal development and peaks around the late–mid fetal stage, concurrent with the onset of synaptogenesis in the cortical plate<sup>14</sup>. To confirm these findings, we performed FISH analysis of *OSTN* expression in tissue sections from a human neocortex at postconception week (pcw) 16 and found that *OSTN* was specifically enriched in the cortical plate of developing neocortex (Fig. 1p, q and Extended Data Fig. 6h). We also compared *OSTN* to *BDNF*, a well characterized activity-regulated secreted factor in rodent and human brains<sup>15</sup>, and found that while *BDNF* was expressed in multiple human brain regions, *OSTN* expression was restricted to the neocortex, and, to a lesser extent, the amygdala (Extended Data Fig. 6a–f). Furthermore, *OSTN* expression was particularly enriched in regions of temporal and occipital cortex, and to a lesser extent in parietal and frontal cortical regions (Extended Data Fig. 6g). Together, these observations indicate that *OSTN* is an activity-regulated gene that is expressed in a subset of neocortical excitatory neurons in the prenatal and postnatal human brain.

### *OSTN* has been repurposed in primates

To determine when and how *OSTN* was repurposed during evolution, so that instead of being narrowly expressed in bone and muscle it is also transcribed in the human brain in response to neuronal activity, we sought to characterize the regulatory elements within the *OSTN* gene locus that mediate activity-dependent *OSTN* expression in human excitatory neocortical neurons. The identification of these regulatory elements has the potential to reveal which species, in addition to humans, express *OSTN* in response to neuronal activity. This should facilitate elucidation of the genetic mechanisms that regulate *OSTN* transcription, and serve as a first step for establishing the function of *OSTN* in the human brain.

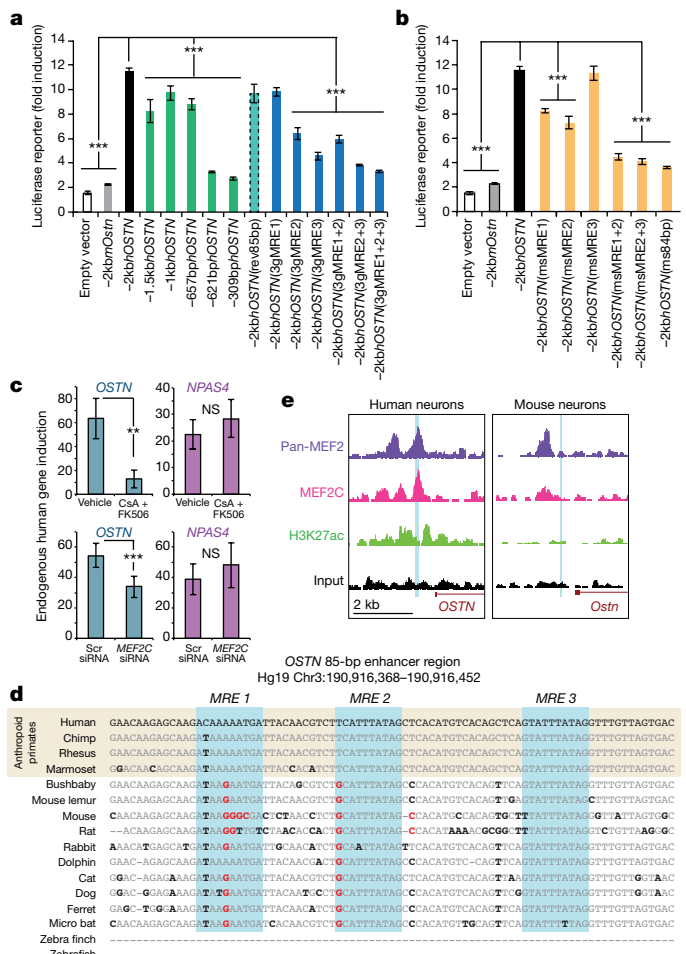


**Figure 1 | *OSTN* is an activity-regulated factor in human neocortex.** **a**, *OSTN* expression in response to membrane depolarization as measured by RNA-seq in hFBCs ( $n = 5$ ), rat ( $n = 3$ ) and mouse ( $n = 4$ ) cortical neuronal cultures. Data shown as mean  $\pm$  s.e.m. **b**, Quantitative RT-PCR for *OSTN* induction in mouse cortical cultures (Ms,  $n = 2$ ), hFBCs (Hum), iCell neurons<sup>35</sup> (Hum. iPS#1,  $n = 3$ ), and *in vitro* differentiated cortical neurons (Hum. iPS#2,  $n = 2$ ) after treatment relative to untreated. KCl depolarization (KCl,  $n = 4$ ), glutamate receptor agonist (NMDA,  $n = 2$ ), calcium chelator (EGTA,  $n = 3$ ), L-type calcium channel blocker nimodopine (Nimod,  $n = 3$ ), NMDA receptor antagonist (2-amino-5-phosphonovaleate; APV,  $n = 2$ ). *OSTN* expression normalized to *GAPDH*.  $n =$  number of biological replicates. Mean  $\pm$  s.e.m., \*\*\* $P < 0.001$ , Student's *t*-test. **c–n**, FISH of depolarized hFBCs showing co-expression of *OSTN* with *MAP2* (pan-neuronal; **c**, **d**), *VGLUT1* (glutamatergic neurons; **e**, **f**), *SATB2* (cortical layer II–IV; **g**, **h**), and *RORB* (layer IV; **i**, **j**). *OSTN* mRNA is excluded from cells positive for *GAD1* (inhibitory neurons; **k**, **l**) and *CTIP2* (cortical layer V; **m**, **n**). Nuclei are marked with DAPI (blue). Scale bar, 6.5  $\mu$ m (**c–f**, **i**, **j**) or 5  $\mu$ m (**g**, **h**, **k–n**). **o**, *OSTN* expression in regions of the developing human brain from HBT database, including striatum (STR), amygdala (AMY), neocortex (NCX), hippocampus (HIP), mediodorsal nucleus of the thalamus (MD), and cerebellar cortex (CBC). **p**, **q**, FISH of human fetal brain (pcw16) coronal section showing selective expression of *OSTN* in the cortical plate of temporal cortex in grayscale (**p**), or overlay (**q**) with DAPI nuclear stain. Scale bar, 390  $\mu$ m. MZ, marginal zone; CP, cortical plate; SP, subplate; IZ, intermediate zone; SVZ, subventricular zone; VZ, ventricular zone.



In principle, the expression of *OSTN* in human but not mouse neurons could be due to *cis*-regulatory element changes in the *OSTN* locus during evolution or to species-specific differences in the expression of a transcription factor(s) that controls *OSTN* expression in neurons, or both. To distinguish between these possibilities, we tested the ability of the 2-kb genomic region directly 5' of the human *OSTN* and mouse *Ostn* transcriptional start sites to drive luciferase reporter gene expression ( $-2\text{kbhOSTN:Fluc}$  and  $-2\text{kbmOstn:Fluc}$ , respectively) in mouse and human neurons. After transfection of these constructs into neurons, we measured the luciferase expression before and after membrane depolarization. We found that  $-2\text{kbhOSTN:Fluc}$  was induced robustly in response to membrane depolarization in mouse or human neurons, whereas the homologous mouse sequence ( $-2\text{kbmOstn:Fluc}$ ) did not drive substantial luciferase expression in neurons of either species (Fig. 2a and Extended Data Fig. 7a, b). These findings suggest that the human specificity of *OSTN* induction is conferred at least in part by the presence of DNA sequences within the human *OSTN* regulatory region that are not present in the homologous mouse sequence. To identify these sequences, we introduced a series of truncations (Fig. 2a, green) and deletions (Extended Data Fig. 8, purple) into the  $-2\text{kbhOSTN}$  regulatory region of the luciferase reporter construct and assessed their effect on the induction of luciferase expression. This analysis identified an 85-bp sequence, located about 600 bp upstream of the site of initiation of *OSTN* mRNA synthesis, that is required for efficient activity-dependent induction of the *OSTN* reporter gene (Fig. 2a, d). The 85-bp sequence induces expression regardless of its orientation in the reporter construct, suggesting that it is a proximal enhancer element (Fig. 2a (teal dashed)). Examination of this sequence revealed three putative myocyte enhancer factor 2 (MEF2)-responsive elements (MREs), referred to here as MRE1–3, two of which were present in human *OSTN* but not the mouse *Ostn* gene (Fig. 2d). These differences could explain why activity induces *OSTN* transcription in human but not mouse neurons.

On the basis of the similarity of the MREs in the human *OSTN* enhancer to the consensus binding site for MEF2 (YTAWWWWTAR)<sup>16,17</sup>, each of these elements was predicted to bind the MEF2 family of four transcription factors, three of which (MEF2A, C, and D) are prominently expressed in the human cortex<sup>13</sup> (<http://www.brainspan.org/>). MEF2 family members are activated by membrane depolarization-induced calcium influx into neurons, which then triggers activation of the calcium-dependent phosphatase calcineurin and the subsequent dephosphorylation and activation of MEF2 (refs 7, 18, 19). To determine whether MEF2 binding to the 85-bp enhancer region is required for the membrane depolarization-dependent induction of reporter gene expression, we either introduced inactivating point mutations<sup>18</sup> (3g mutations) into the three MREs of the 85-bp enhancer (Fig. 2a, blue), or exposed neurons to two small-molecule calcineurin inhibitors (CsA and FK506; Extended Data Fig. 7c), which inhibit membrane depolarization-dependent dephosphorylation and activation of MEF2. These experiments demonstrated that activity-dependent induction of  $-2\text{kbhOSTN:Fluc}$  was calcineurin-dependent and required intact MRE sequences within the 85-bp enhancer. Furthermore, chromatin immunoprecipitation sequencing (ChIP-seq) for MEF2 family members showed enrichment in human but not mouse neurons at the 85-bp enhancer region (Fig. 2e and Extended Data Fig. 7d), indicating that MEF2 binds to the endogenous *OSTN* locus in human neurons. It is noteworthy that an additional site of shared MEF2 binding was observed in human and mouse neurons upstream of the 85-bp enhancer; in mouse neurons, this site lacks epigenomic marks associated with active chromatin (for example, no enrichment of H3K27ac in Fig. 2e and Extended Data Fig. 7e) and thus is likely to represent a transcriptionally inert site of MEF2 binding in mice. Our findings suggest that evolution of the tandem MEF2-binding sites within the human 85-bp enhancer element probably explains, at least in part, why *OSTN* is induced in response to neuronal activity in human but not mouse neurons.



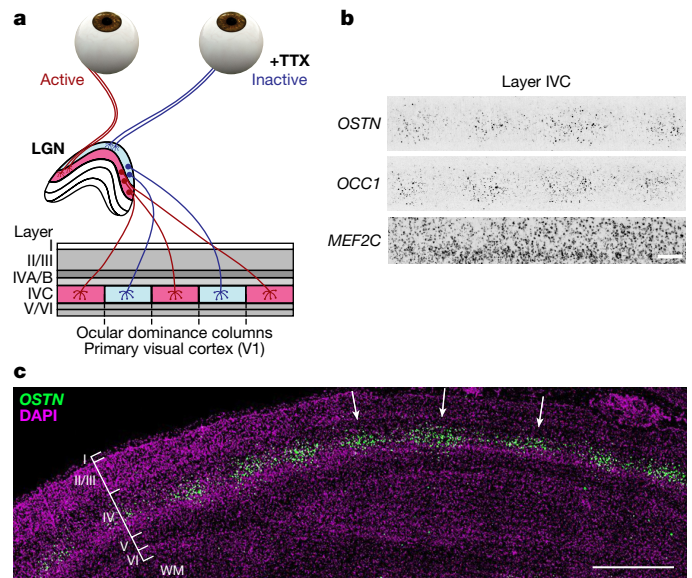
**Figure 2 | Primate-specific enhancer regulation by MEF2 drives neuronal activity-dependent induction of *OSTN*.** **a**, The 2-kb human genomic region upstream of *OSTN* ( $-2\text{kbhOSTN}$ , black), but not the homologous mouse sequence ( $-2\text{kbmOstn}$ , grey), drives the expression of a luciferase reporter gene in response to membrane depolarization (55 mM KCl) in mouse cortical cultures. Truncations (green) of  $-2\text{kbhOSTN}$  identify an 85-bp region that is critical for luciferase induction. Reversal of the 85-bp sequence orientation does not affect gene induction (teal dashed), unlike point mutations that disrupt MEF2 binding (3g mutations, blue). **b**, Luciferase assays with  $-2\text{kbhOSTN}$ , mutating the individual MREs or the entire 85-bp enhancer to the corresponding mouse sequence (gold). See Extended Data Fig. 8b. One-way ANOVA degrees of freedom = (33, 266),  $P < 1.0 \times 10^{-13}$ , Holm-Sidak overall error rate of 0.05, \*\*\* $P < 1.22 \times 10^{-5}$ . **c**, Quantitative RT-PCR for endogenous *OSTN* and *NPAS4* gene induction in days *in vitro* (DIV) 21 hFBCs following 6 h of KCl depolarization. Addition of calcineurin inhibitors (CsA and FK506,  $n = 4$ ) or siRNAs targeting *MEF2C* (*MEF2C* siRNA,  $n = 4$ ) specifically affected *OSTN* induction.  $n$  = number of biological replicates. Mean  $\pm$  s.e.m., Student's *t*-test \*\*\* $P < 0.01$ , \*\*\* $P < 0.001$ , NS, not significant. **d**, Sequence conservation map of the 85-bp enhancer element. Conserved bases are grey, non-conserved bases are black, gaps in alignment are dashed, and sequence changes predicted to disrupt MEF2 binding are red. **e**, ChIP-seq using a pan-MEF2 antibody (purple), a MEF2C-specific antibody (fuschia), and an H3K27ac antibody (green) indicating active chromatin from hFBCs (left) and mouse cortical cultures (right) shows enrichment for MEF2 binding at the primate-specific 85-bp enhancer region (blue highlight) in hFBCs but not in mouse neurons.

The identification of regulatory sequences within the human *OSTN* gene that confer the response to neuronal activity allowed us to investigate the evolutionary origin of activity-dependent *OSTN* expression in the brains of other mammals (for example, macaque, marmoset, mouse, rabbit and cat). We interrogated available genome sequences for the

presence of the identified MRE sites within the regulatory region of the *OSTN* gene. This analysis revealed that the combined MRE1/2 motifs are conserved specifically in anthropoid primates but are not present in prosimian primates, rodents or a variety of other species examined (Fig. 2d). Moreover, replacing these two MREs, or the entire 85-bp enhancer element of the human *OSTN* sequence, with the corresponding mouse sequences, disrupted depolarization-induced  $-2\text{kbhOSTN}$  reporter gene expression in neurons (Fig. 2b (gold)). Finally, the activity-dependent increase in endogenous *OSTN* transcript in human neurons was specifically impaired by calcineurin inhibitors (CsA and FK506) or a small inhibitory RNA (siRNA) against MEF2C (per cent knockdown,  $48.96 \pm 5.7$ ), the most abundant MEF2 family member in the human neocortex (Fig. 2c). Together, these findings indicate that the repurposing of *OSTN* as a neuronal activity-responsive gene occurred during the evolution of the anthropoid primates through subtle sequence changes (around 5–10 nucleotides) within a highly conserved region of the *OSTN* gene, and that these changes led to the emergence of a new activity-regulated enhancer element with functional MEF2 sites.

### *OSTN* is regulated by activity in vivo

The finding that the 85-bp *OSTN* enhancer (including the MREs) is conserved in primates provided us with the opportunity to test whether the induction of *OSTN* that we observe in cultured neurons in response to membrane depolarization occurs in the intact brain in response to sensory stimuli. If so, it should be possible to use FISH to identify the specific neurons in the cortex that express activity-dependent *OSTN* mRNA in order to gain insight into the function of *OSTN* in the brain. Towards this end, we used a monocular inactivation assay in the Old World Macaque monkey (*Macaca mulatta*), an established model for studying activity-dependent cortical plasticity and gene expression in the primary visual cortex<sup>20,21</sup>. Vision is the dominant sensory modality in primates, and during evolution the size and complexity of the visual cortex in primates have expanded more than those of any other species in the mammalian lineage<sup>22</sup>. In primates, under normal conditions, inputs from the left and right eyes remain segregated through the lateral geniculate nucleus (LGN) of the thalamus and innervate the geniculorecipient layers of the primary visual cortex (area 17 or V1). In V1, the two eyes' inputs form alternating stripes termed ocular dominance columns (ODCs)<sup>23</sup> (Fig. 3a). Monocular inactivation in adult monkeys results in a characteristic pattern of alternating active and inactive ODCs that can be used to assess sensory activity-dependent gene expression when labeled by FISH<sup>21</sup>. To test whether *OSTN* expression in the macaque brain is induced by visual experience, we blocked retinal activity in one eye of adult rhesus macaques using tetrodotoxin injection and one day later removed, froze and sectioned the occipital cortex. FISH analysis in sections of V1 revealed that, after monocular inactivation, *OCC1* (also known as *FSTL1*), a known activity-dependent gene in macaque V1 (ref. 24), was expressed in an alternating pattern of greater and lesser gene expression (Fig. 3b). Thus, *OCC1* serves as a marker of ODCs receiving input from the active eye. *OSTN* mRNA was found to be expressed at a higher level in active than in inactive ODCs, indicating that *OSTN* mRNA expression in the primate brain is induced by sensory experience (Fig. 3b, c and Extended Data Fig. 9a–c). Nuclear staining and FISH with probes that hybridize to mRNAs that are specifically expressed in glutamatergic neurons (*VGLUT1*<sup>+</sup>) (Extended Data Fig. 9e) or cortical layer IV (*RORB*<sup>+</sup>) neurons (Extended Data Fig. 9f) revealed that *OSTN* expression was almost exclusively restricted to excitatory neurons in cortical layer IVC, the neurons that receive direct input from the LGN<sup>23</sup>. In addition, *OSTN* was coexpressed with *MEF2C* and *MEF2A* in layer IVC neurons (Extended Data Fig. 9g, h), consistent with the finding that in primates the activity-dependent induction of *OSTN* is mediated by MEF2 activation. Together, these observations suggest that activity- and MEF2-dependent regulation of *OSTN* expression has evolved in the primate brain to regulate excitatory neuron function specifically in the



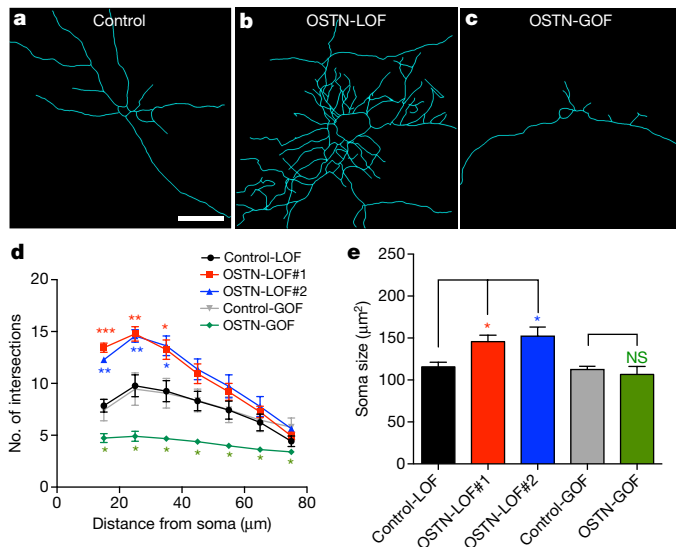
**Figure 3 | *OSTN* expression is induced by sensory experience in the primate cortex.** **a**, Schematic diagram of macaque visual pathway and monocular inactivation assay. For simplicity, only one hemisphere is depicted. **b**, High magnification FISH images of monocularly inactivated layer IVC neurons for *OSTN*, *OCC1*, and *MEF2C* transcripts. Scale bar, 188  $\mu\text{m}$ . **c**, FISH image of *OSTN* from a radial section of monocularly inactivated macaque primary visual cortex. Scale bar, 1,000  $\mu\text{m}$ . Cortical layers I–VI and white matter (WM) are labelled, and active ocular dominance columns are indicated with arrows.

input layers of the larger and more complex neocortex of primates. In line with this idea, we found that *OSTN* expression was also enriched in layer IV of the macaque multimodal parietal cortex (Extended Data Fig. 9d). While *OSTN* expression is regulated by the activity of afferent inputs in the adult, it may also be driven by spontaneous activity during the early stages of cortical development<sup>25,26</sup> before the majority of synaptic inputs are established.

### *OSTN* regulates dendritic growth

To investigate the function of *OSTN* in the primate brain, we adapted a chronic activity culture paradigm<sup>27</sup> (CAP) in which neurons were maintained under mildly depolarizing conditions (20 mM KCl) to mimic neuronal activity, drive *OSTN* transcription and secretion (Extended Data Fig. 10a, b), and elicit activity-driven changes in neuronal morphology and function. Under these culture conditions, *OSTN* induction was suppressed by the addition of either of two different *OSTN*-specific siRNAs (84% and 94% knockdown; Extended Data Fig. 10b) or enhanced by overexpression of *OSTN*, and the effect on dendritic growth was assessed. Introduction of a GFP-based reporter driven by the  $-2\text{kbhOSTN}$  regulatory region ( $-2\text{kbhOSTN}:\text{GFP}$ ) allowed us to visualize *OSTN*-expressing cells (Extended Data Fig. 10c, d). Following *OSTN* knockdown (loss-of-function, LOF) or overexpression (gain-of-function, GOF), cultures were grown for three days to allow the accumulation of GFP and *OSTN* and then immunostained for GFP and MAP2 to visualize the somata and dendritic arbors of *OSTN*-expressing cells. Sholl analyses revealed that reductions in *OSTN* substantially increased dendritic branch number and complexity, whereas increasing *OSTN* levels led to marked inhibition of dendritic branching (Fig. 4a–d). These findings suggest that *OSTN* functions to restrict dendritic growth in response to membrane depolarization. Notably, neuronal soma size was also increased when *OSTN* expression was reduced, suggesting that *OSTN* may restrict neuronal growth overall (Fig. 4e). Furthermore, immunostaining for endogenous *OSTN* protein in hFBCs indicated that *OSTN* was localized to the soma and primary dendrites of neurons, consistent





**Figure 4 | OSTN regulates activity-dependent dendritic growth.**

**a–c**, Representative GFP tracings of individual hFBC neurons transfected with  $-2\text{kbhOSTN:GFP}$  and treated with scrambled siRNA (Control, **a**) or OSTN siRNA#1 (OSTN-LOF, **b**), or co-transfected with an OSTN overexpression construct (OSTN-GOF, **c**) in the presence of CAP for 3 days. Scale bar, 30  $\mu\text{m}$ . **d**, Sholl analysis plot showing numbers of dendritic intersections as a function of distance from soma for scrambled siRNA (Control-LOF, in 5 biological replicates, 184 neurons analysed (5; 184)), OSTN siRNA#1 (OSTN-LOF#1 (5; 166)), OSTN siRNA#2 (OSTN-LOF#2 (3; 104)), empty vector transfected (Control-GOF (4; 99)), and OSTN overexpression construct transfected (OSTN-GOF (4; 112)). **e**, Summary of average soma sizes. Mean  $\pm$  s.e.m., \*\*\* $P < 0.001$ ; \*\* $P < 0.01$ ; \* $P < 0.05$ , NS, not significant; Student's  $t$ -test.

with the proposed role of OSTN in dendritic growth and arborization (Extended Data Fig. 10e–g).

## Conclusions

In this study, we have identified an example of the evolutionary repurposing of a secreted factor via genetic changes that give rise to a new primate-specific enhancer sequence that binds MEF2. MEF2 has been implicated in human cognition because *MEF2C* mutations resulting in haploinsufficiency represent a relatively common cause of intellectual disability in humans<sup>28,29</sup>. In addition, MEF2A- and MEF2C-binding sites are enriched in genes associated with idiopathic autism spectrum disorder (ASD), supporting the idea that dysregulation of activity-dependent transcription plays a role in the pathogenesis of ASD<sup>30</sup>. One intriguing possibility is that OSTN is a critical target of MEF2 whose function is disrupted in MEF2-dependent brain disorders.

Our findings indicate that MEF2, by regulating OSTN expression, restricts the dendritic growth of layer IVC neurons and thus may play an important role in the organization of dense primate neocortical networks<sup>31–33</sup>. However, the potential contribution of OSTN function to primate cognitive processing and what drove OSTN repurposing specifically in primates during evolution are still unknown. Notably, evolution of the proximal enhancer required for efficient induction of OSTN by neuronal activity occurred selectively in anthropoid primates (monkeys, apes and humans), which are characterized by diurnal lifestyles, large neocortical volumes and complex socio-cognitive skills.

As neuronal activity has been shown to trigger distinct transcriptional responses in different neuronal subtypes<sup>34</sup>, it seems likely that, in addition to OSTN, other primate-specific neuronal factors remain to be discovered. In this study, we have begun to uncover evolved primate-specific gene regulatory changes that alter intrinsic cellular

mechanisms associated with activity-driven plasticity in the neocortex. Investigation of additional neural factors and their regulatory elements may provide further insights into important aspects of human brain development, function and disease.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 January; accepted 19 September 2016.

- Defelipe, J. The evolution of the brain, the human nature of cortical circuits, and intellectual creativity. *Front. Neuroanat.* **5**, 29 (2011).
- Buñill, E., Agustí, J. & Blesa, R. Human neoteny revisited: The case of synaptic plasticity. *Am. J. Hum. Biol.* **23**, 729–739 (2011).
- Geschwind, D. H. & Rakic, P. Cortical evolution: judge the brain by its cover. *Neuron* **80**, 633–647 (2013).
- Johnson, M. B. et al. Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat. Neurosci.* **18**, 637–646 (2015).
- Pollen, A. A. et al. Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015).
- Kwan, K. Y. et al. Species-dependent posttranscriptional regulation of NOS1 by FMRP in the developing cerebral cortex. *Cell* **149**, 899–911 (2012).
- Flavell, S. W. et al. Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection. *Neuron* **60**, 1022–1038 (2008).
- Kim, T.-K. K. et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187 (2010).
- Greer, P. L. & Greenberg, M. E. From synapse to nucleus: calcium-dependent gene transcription in the control of synapse development and function. *Neuron* **59**, 846–860 (2008).
- Banzet, S. et al. Musclin gene expression is strongly related to fast-glycolytic phenotype. *Biochem. Biophys. Res. Commun.* **353**, 713–718 (2007).
- Subbotina, E. et al. Musclin is an activity-stimulated myokine that enhances physical endurance. *Proc. Natl Acad. Sci. USA* **112**, 16042–16047 (2015).
- Moffatt, P. et al. Osteocrin is a specific ligand of the natriuretic peptide clearance receptor that modulates bone growth. *J. Biol. Chem.* **282**, 36454–36462 (2007).
- Kang, H. J. et al. Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).
- Molliver, M. E., Kostović, I. & van der Loos, H. The development of synapses in cerebral cortex of the human fetus. *Brain Res.* **50**, 403–407 (1973).
- Hong, E. J., McCord, A. E. & Greenberg, M. E. A biological function for the neuronal activity-dependent component of Bdnf transcription in the development of cortical inhibition. *Neuron* **60**, 610–624 (2008).
- Gossett, L. A., Kelvin, D. J., Sternberg, E. A. & Olson, E. N. A new myocyte-specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. *Mol. Cell. Biol.* **9**, 5022–5033 (1989).
- Andzel, M. M. et al. MEF2D drives photoreceptor development through a genome-wide competition for tissue-specific enhancers. *Neuron* **86**, 247–263 (2015).
- Flavell, S. W. et al. Activity-dependent regulation of MEF2 transcription factors suppresses excitatory synapse number. *Science* **311**, 1008–1012 (2006).
- Potthoff, M. J. & Olson, E. N. MEF2: a central regulator of diverse developmental programs. *Development* **134**, 4131–4140 (2007).
- Horton, J. C. & Hocking, D. R. Monocular core zones and binocular border strips in primate striate cortex revealed by the contrasting effects of enucleation, eyelid suture, and retinal laser lesions on cytochrome oxidase activity. *J. Neurosci.* **18**, 5433–5455 (1998).
- Takahata, T., Higo, N., Kaas, J. H. & Yamamori, T. Expression of immediate-early genes reveals functional compartments within ocular dominance columns after brief monocular inactivation. *Proc. Natl Acad. Sci. USA* **106**, 12151–12155 (2009).
- Kaas, J. H. The evolution of neocortex in primates. *Prog. Brain Res.* **195**, 91–102 (2012).
- Kolb, H., Fernandez, E. & Nelson, R. *Webvision: The Organization of the Retina and Visual System* (University of Utah Health Sciences Center, 1995).
- Tochitani, S., Liang, F., Watakabe, A., Hashikawa, T. & Yamamori, T. The occ1 gene is preferentially expressed in the primary visual cortex in an activity-dependent manner: a pattern of gene expression related to the cytoarchitectonic area in adult macaque neocortex. *Eur. J. Neurosci.* **13**, 297–307 (2001).
- Kostović, I. & Judoš, M. The development of the subplate and thalamocortical connections in the human foetal brain. *Acta Paediatr.* **99**, 1119–1127 (2010).
- Moore, A. R. et al. Connexin hemichannels contribute to spontaneous electrical activity in the human fetal cortex. *Proc. Natl Acad. Sci. USA* **111**, E3919–E3928 (2014).
- O'Leary, T., van Rossum, M. C. & Wyllie, D. J. Homeostasis of intrinsic excitability in hippocampal neurones: dynamics and mechanism of the response to chronic depolarization. *J. Physiol. (Lond.)* **588**, 157–170 (2010).
- Zweier, M. et al. Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. *Hum. Mutat.* **31**, 722–733 (2010).



29. Bienvenu, T., Diebold, B., Chelly, J. & Isidor, B. Refining the phenotype associated with MEF2C point mutations. *Neurogenetics* **14**, 71–75 (2013).
30. Parikshak, N. N. *et al.* Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**, 1008–1021 (2013).
31. Lund, J. S., Holbach, S. M. & Chung, W. W. Postnatal development of thalamic recipient neurons in the monkey striate cortex: II. Influence of afferent driving on spine acquisition and dendritic growth of layer 4C spiny stellate neurons. *J. Comp. Neurol.* **309**, 129–140 (1991).
32. Herculano-Houzel, S. Neuronal scaling rules for primate brains: the primate advantage. *Prog. Brain Res.* **195**, 325–340 (2012).
33. Srinivasan, S., Carlo, C. N. & Stevens, C. F. Predicting visual acuity from the structure of visual cortex. *Proc. Natl Acad. Sci. USA* **112**, 7815–7820 (2015).
34. Mardinly, A. R. *et al.* Sensory experience regulates cortical inhibition by inducing IGF1 in VIP neurons. *Nature* **531**, 371–375 (2016).
35. Berry, B. J. *et al.* Morphological and functional characterization of human induced pluripotent stem cell-derived neurons (iCell Neurons) in defined culture systems. *Biotechnol. Prog.* **31**, 1613–1622 (2015).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank E. Curran, T. Hartmann, P. Schade, P. Zhang, J. M. Gray, M. Hemberg, X. Adiconis, J. Z. Levin, J. Zieg, D. R. Hochbaum, T. J. Cherry and M. M. Andzelm for their technical assistance or advice. This work was supported by grants from the NIH: P50MH106933 and 1RC2MH089952

(M.E.G.), 5F32NS086270 (G.L.B.), EY16187 (M.S.L.), EY12196 (V.K.B.), and T32GM007753 (A.N.M. and E.D.). B.A. is supported by The Ellen R. and Melvin J. Gordon Center for the Cure and Treatment of Paralysis.

**Author Contributions** B.A., G.L.B., and M.E.G. performed or directed all experiments and wrote the manuscript with E.C.G. B.A., D.A.H. and M.G.Y. performed or analysed gene expression experiments. G.L.B. performed iPSC differentiation and luciferase reporter assays with assistance from M.G.Y., K.M. and M.B.-S. V.K.B., G.L.B. and M.S.L. performed monocular inactivation experiments. B.A., M.G.Y. and G.L.B. performed FISH. B.A. and M.G.Y. performed dendritic growth assays. E.-L.Y. and N.Sh. contributed to overexpression studies. A.N.M. and A.A.R. cloned initial reporter constructs. I.S. provided RiboTag-Seq data and mouse brain sections. L.S.H. generated the OSTN antibody. M.P. and N.Se. provided human brain sections. M.C., J.N.P. and C.A.W. provided human tissue for initial culture experiments. M.A., B.A., M.G.Y., and E.D. performed ChIP experiments. C.R.S. assisted with RNA-seq experiments.

**Author Information** Raw and processed RNA-seq data have been submitted to the NCBI Gene Expression Omnibus under accession number GSE78688. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.E.G. ([meg@hms.harvard.edu](mailto:meg@hms.harvard.edu)).

**Reviewer Information** *Nature* thanks M. Oldham, F. Polleux and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were blinded to sample identity during imaging and quantification for Sholl analysis.

**Human neuronal cultures.** Initial research performed on samples of human origin was conducted with the approval of the Beth Israel Deaconess Medical Center Committee on Clinical Investigations (IRB#2001P000527). Fetal brain tissue was received after release from clinical pathology, with a maximum post-mortem interval of 4 h. Cases with known pathology were excluded. Tissue was transported in HBSS medium (Life Technologies, Carlsbad, CA) on ice to the laboratory for processing. For the initial fetal brain culture (hFBC) experiments, obtained fetal tissue was dissected and trypsinized. The resulting cell suspensions were gradient purified to remove cell debris using OptiPrep (Sigma; St. Louis, MO) density gradients, adapting the published protocol<sup>36</sup>. Subsequent studies used primary human neuronal cells purchased from ScienCell (Carlsbad, CA).

**KCl depolarization of neurons.** For acute KCl depolarization of neurons, neuronal cultures were first silenced overnight in culture medium with 1  $\mu$ M tetrodotoxin (TTX) and 100  $\mu$ M D-APV (both Tocris). The next day, samples were incubated for 0, 1 or 6 h in 55 mM KCl before collection (3  $\times$  KCl solution: 170 mM KCl, 10 mM HEPES, 2 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, pH 7.4). All cultures were monitored during treatment and no adverse effects to cell health were observed. Additionally, we did not observe upregulation of the excitotoxicity-induced gene *Clca1* (data not shown). Samples in Fig. 1b were harvested without stimulation as well as 6 h after KCl (55 mM) depolarization or glutamate receptor agonist (NMDA, 20  $\mu$ M) treatment. Prior to stimulation, samples were treated with the calcium chelator EGTA (5 mM), the L-type channel blocker nimodopine (Nimod, 5  $\mu$ M), or the NMDA antagonist APV (100  $\mu$ M) as indicated.

**Transcript sequencing and RNA-seq analysis.** Previous studies detected robust changes in expression of activity-dependent genes within KCl membrane-depolarized neuronal cultures using three or fewer biological replicates<sup>7,8</sup>. For RNA-seq experiments, hFBCs from 5 different individuals (ScienCell) were grown according to the supplier's instructions. At DIV15, neurons were silenced overnight and then KCl depolarized (see above) for 0, 1 or 6 h before collection. Total RNA was isolated from cultures using Trizol (Invitrogen). After DNase treatment and rRNA depletion, strand-specific and paired-end cDNA libraries were generated using the PE RNA-seq library kit (Illumina). Ribosomal RNA depletion was performed using the RiboMinus Eukaryote Kit for RNA-seq (ThermoFisher) and verified using a Bioanalyzer RNA Nano kit (Agilent). Fragment ends were sequenced to produce strand-specific paired-end reads of at least 76 base pairs (bp) in length. RNA-seq was performed using HiSeq 2000 at the Broad Institute and BGI. For each sample at each stimulation time point, both sets of single-end reads were separately aligned to the human genome (hg19 assembly) using the Burrows-Wheeler Aligner (BWA) program, allowing for up to five mismatches. In addition to the usual 24 chromosomal targets, a set of  $\sim$ 7 million short splice-junction sequences (see below) were also provided as targets and incorporated into the BWA index. Each sequenced library comprised 47–240 million reads, of which 55–95% were mappable. Of those reads that did map at all, typically  $\sim$ 90% were aligned uniquely. We found that for our purposes (expression levels and UCSC Genome Browser tracks) full RNA fragment reconstruction was not necessary, so going forward our data sets comprised just the uniquely mapped single-end reads from 'end #1'.

The splice-junction target sequences were based on the NCBI RefSeq database for human 37.1 (hg19). For each annotated transcript, we noted all subsets of two or more exons, not necessarily adjacent, that could be spliced together to produce a sequence at least as long as the read length (76 bp). Each of these sequences was then trimmed to the maximum number of bases such that a read mapping to the sequence would have to cross only these ordered exons' splice junction(s). This procedure produced a library of all unique sets of exons whose intragenic splice junctions could possibly be covered by a read of the given length, based on the RefSeq annotation of exonic loci. Aligned reads thus had the opportunity to align either to genomic (chromosomal) sequences or to exon-junction-crossing sequences found only in mature mRNA. Multiple reads whose 5' ends were assigned to the same locus on the same strand were not flattened to a single count.

An in-house software tool, MAPtoFeatures<sup>37</sup>, was used to quantify expression levels for individual genes as follows. A database of genic features (coding sequences (CDSs) and untranslated regions) was constructed from all 29,149 transcripts annotated in RefSeq (human 37.1, 12 March 2009). Merged genes were constructed by combining all exons in all transcripts assigned to each distinct gene; the resulting segments defined the gene's exonic coordinates used here (with the gaps between them defining introns). Genes with zero CDS exons were labelled 'non-coding'. These 20,066 genes were supplemented with 1,723 additional noncoding genes specified by the loci of all ribosomal RNA genes obtained from RepeatMasker

(where the options Variations and Repeats, rmsk.repFamily = "rRNA" yielded 391 LSU-rRNA\_Hsa; 71 SSU-rRNA\_Hsa; 1,261 5S). The purpose of this step was to filter out reads stemming from transcription of repeats and rRNA genes, which tend to get populated to inconsistent degrees from sample to sample depending on the variable quality of rRNA depletion.

Reads that aligned uniquely were then queried for their intersection with the exonic ranges of any of the above 21,789 genes, including exon-exon splice junctions. The total number of read bases that overlapped an exonic range in the sense direction was divided by the range's length to give an average exonic read density (that is, coverage). All reads were assigned to genes or to intergenic regions. However, only those reads not assigned to noncoding genes counted towards the total normalization count *N*, which ultimately afforded a more stable comparison of expression levels between samples than simply using the total number of reads. All read densities were normalized to a reference total of 10 million reads and a reference read length of 35 bp through multiplication by the factor  $(10^7/N) \times (35 \text{ bp}/76 \text{ bp})$ . Division of these normalized densities by 0.35 yielded expression levels in alternative units of reads per kilobase of transcript per million mapped reads (RPKM).

Differential expression analyses aimed to produce fold change ratios (between 0 and 1 h or 0 and 6 h time points) and their statistical significance for every expressed gene. For this purpose read counts were preferable to read densities for two reasons: sample-independent parameters such as a gene's exonic length cancel out of such ratios, and a null model of low read counts would in any case require discrete data. Whole-read counts were monitored for each sample for reads that fell entirely within single exons as well as for those that crossed exon-exon splice-junction boundaries captured by the aforementioned splice library, a not insubstantial fraction (10–30%) of all exonic reads. Relatively few reads (<1–3%) crossed exon-intron boundaries; nevertheless, the exonic fraction of such reads was counted towards a gene's total 'fractional read counts' (frds), rounded up to the nearest integer. Genes were further processed for differential expression between 0 h and 1 h (or 6 h) depolarization only if they passed a minimal read counts filter and an expression level background filter. All five replicates needed to have a total of at least three fractional reads per gene at each time point. Furthermore, the geometric mean density over all non-zero density values (up to five replicates) needed to be at least 0.20 (RPKM  $\geq$  0.5714) at either 0 h or the later time point. The filtered tables of frds ( $\sim$ 12,000 genes over 5 samples per time point) was then taken as input to the R Bioconductor package edgeR<sup>38</sup>. This tool was appropriate for our samples because it is able to model low counts subject to biological variability via negative binomial distributions. Normalization factors were calculated using the default TMM method; dispersion was estimated 'tagwise'. In order to control the false discovery rate (FDR), *P* values were adjusted via the usual Benjamini-Hochberg (BH) procedure<sup>39</sup>.

RNA-seq was performed on cultured mouse and rat cortical cells from E16.5 C57BL/6 wild-type mice and E17 Long-Evans rats as described<sup>7,8</sup>. Three biological replicates of mouse and rat cultured neurons were KCl depolarized (see above) for 1 or 6 h at DIV7. Paired-end reads of length 76 bp were sequenced on an Illumina platform and aligned to the mouse genome (GRCm38/mm10 assembly, Dec. 2011) as described above for the human samples. For mouse, the 95,023 transcripts annotated in RefSeq and assigned to 33,102 genes in mm10 were supplemented with 1,563 additional rRNA genes from RepeatMasker; a library of splice-junction target sequences was constructed; and normalized RPKM expression levels were assigned to merged genes as described above. Similarly for rat, 58,438 transcripts annotated in RefSeq (RGSC 6.0/rn6, July 2014) and assigned to 28,022 genes were supplemented with 1,641 rRNA genes from RepeatMasker; a splice library was constructed; and normalized RPKM levels were assigned to merged genes.

The RiboTag-Seq experiments were performed as described<sup>34,40</sup>. Briefly, double-heterozygous mice (Cre/+, RiboTag/+) were reared under a standard light cycle and then housed in constant darkness for two weeks starting from P42; at P56, all mice were either euthanized in the dark (dark-housed condition) or light-exposed for 1, 3, or 7.5 h before being euthanized. Visual cortices were dissected, and immunoprecipitation and purification of ribosome-associated RNA was performed as described<sup>34</sup>. Visual cortices from three individual animals (each sample contained both male and female animals) were pooled for each biological replicate, and three biological replicates were performed. For all RNA samples of sufficient integrity (RIN >8.0), 5–10 ng RNA was SPRI-amplified with the Ovation RNA Amplification System V2 (NuGEN). For RNA-seq analyses, 2 mg of each amplified cDNA was fragmented to a length of 200–400 bp and used for Seq library preparation using the PrepX DNA kit (IntergenX). The completed libraries were sequenced on an Illumina HiSeq 2000 instrument following the manufacturer's standard protocols for single-end 50-bp sequencing with single index reads, and reads were mapped to the mouse genome (NCBI37/mm9 assembly, July 2007) using TopHat (v2.0.13) and Bowtie (2.1.0.0).

**Data transformation and filters.** The scatterplots in Extended Data Fig. 3a, b show expression levels of data for five samples versus five samples, for either 1 h or 6 h post-KCl-stimulation versus unstimulated (0 h). At each time point the five data sets were log-transformed and quantile normalized, with only non-zero data included. Each point shown in the figures represents a gene with at least one non-zero value at each of the two time points; at each time point error bars show  $\pm 1$  s.e. over  $\log_{10}$  expression levels while their crossing point lies at the two-dimensional mean of the gene's  $\log_{10}$  values (that is, at the geometric mean).

The log of the ratio of mean values (fold-change) is proportional to the distance from the main diagonal. Off-diagonal grey lines mark up or down ratios of 10 and 100. Genes with the most unusual fold-changes were filtered by magnitude, significance, and comparison to background expression levels. A background threshold was chosen at the density value 0.20 (that is,  $\text{RPKM} = 0.20/0.35 = 0.5714 \equiv B$ ). Those genes with a mean density exceeding this threshold at either 0 h or the later time point were submitted to the R package edgeR, as described above, for evaluation of the significance of their differential expression, including BH-adjusted  $P$  values to control FDR. Highlighted points have adjusted  $P$  consistent with an FDR threshold 15% or less and fold changes either  $\geq 2.0$  or  $\leq 0.5$ . For 1 h versus 0 h, 9 of 17,323 genes shown passed all these filters (all upregulated); for 6 h vs. 0 h, 185 of 17,224 did (73 upregulated, 112 downregulated). The highest fold-change of 102.9 $\times$  was ascribed to *OSTN* at 6 h of membrane depolarization; the greatest change at 1 h was 47.1 $\times$  for *NPAS4*. Supplementary Table 1 contains the lists for early-response genes (ERGs) and late-response genes (LRGs). We categorized a gene as an ERG if the edgeR calculated fold-change (FC) at 1 h was greater than the FC at 6 h. If a gene had a greater FC at 6 h, it was categorized as an LRG. Note that fold changes inferred from Extended Data Fig. 3a, b, based on mean expression levels, may differ slightly from the fold change values cited here, which are taken from Supplementary Tables 1 and 4 and are instead based on read counts and their normalization as calculated by edgeR.

The quantile distribution for the five unstimulated samples is plotted in Extended Data Fig. 2a against  $\log_{10}$ -expression level (RPKM units) with a colour scale for the heatmaps.

**Spearman correlation heat map and panel of human tissues.** Spearman correlations ( $r_s$ ) in Extended Data Fig. 1f were calculated using gene expression levels from our five unstimulated hFBC replicates (H1–H5). Only those 11,711 RefSeq genes that were annotated in hg19 as coding genes and that had non-zero expression levels in all samples were included in this calculation. The dendrogram in Extended Data Fig. 1g was based on the hierarchical clustering of the expression levels of these five samples plus ten previously sequenced human tissues<sup>37</sup> (GEO accession number GSE48889), including whole brain, with distance measure 1– $r_s$ . In order to emphasize informative genes for this comparison, a subset of  $\sim 3,000$  coding genes was identified that had non-zero expression in at least 5 of the 10 tissues and could be loosely classified as expression outliers in any one of the 10 tissues (Grubbs's test,  $\alpha = 0.10$ , Bonferroni adjusted). The displayed horizontal ordering minimizes the total of the distance measures between adjacent samples. Brain clusters with H1–H5 while correlations among the remaining tissues are unstructured.

**BrainSpan data.** Expression levels from the publicly available BrainSpan atlas (<http://www.brainspan.org/>) are derived from RNA-seq for 22,327 genes and 578 human samples, including 41 individuals, 30 different ages ranging from embryonic to adulthood and 26 specific brain regions. The data shown for *OSTN* and *BDNF* in Extended Data Fig. 6a–f cover five separate regions: amygdaloid complex, AMY; cerebellar cortex, CBC; hippocampus, HIP; mediodorsal nucleus of thalamus, MD; and striatum, STR. We further combined the data for 11 cortical regions under 'Neocortex' (NCX): primary auditory cortex (core), A1C; dorsolateral prefrontal cortex, DFC; posteroventral (inferior) parietal cortex, IPC; inferolateral temporal cortex, ITC; primary motor cortex, M1C; anterior (rostral) cingulate (medial prefrontal) cortex, MFC; orbital frontal cortex, OFC; primary somatosensory cortex, S1C; posterior (caudal) superior temporal cortex, STC; primary visual cortex (striate cortex), V1C; and ventrolateral prefrontal cortex, VFC. *OSTN* expression data from these neocortical regions were also grouped into four anatomical categories and displayed in Extended Data Fig. 6g. For each gene, data for all samples at all available time points in each brain region were fit via a local polynomial regression (the Loess function in R version 3.0.2) and shown as mean Loess curves interpolated across the whole age range. The width of the one-standard-error side bands shown in Extended Data Fig. 6a–g were similarly calculated via a Loess fit to the standard errors deduced at whatever ages data were available in each region for each gene.

**Quantitative PCR.** Total RNA was isolated from human and mouse neuronal cultures using Trizol. Isolated RNA was treated with Amplification Grade DNaseI (Invitrogen), and cDNA libraries were synthesized using a High Capacity cDNA

Reverse Transcription Kit (Applied Biosystems). The cDNA was the source of input for quantitative RT-PCR, using a Step One Plus Real-Time PCR Instrument and SYBR Green reagents (Applied Biosystems). The relative expression plot was constructed using concentration values that were normalized to corresponding *GAPDH* concentrations. The following primer sets were used for qPCR experiments: Human *OSTN* F-5'-CAGGAAAAGTCTCTCAGTAGATG-3', R-5'-GCAAGAGTTTGTCTGTCAAGTCA-3'; Mouse *Ostn* F-5'-CCATGGATCGGATTGGTAGA-3', R-5'-GCCATCTCACACAAGTAAGTCG-3'; Human *NPAS4* F-5'-TGGGTTTACTGATGAGTTGCAT-3', R-5'-TCCCCTCCACTTCCATCTT-3'; Human *GAPDH* F-5'-GTCTCTCTGACTTCAACAGCG-3', R-5'-ACCACCCTGTTGCTGTAGCCAA-3'; Mouse *Gapdh* F-5'-CATCACTGCCACCAAGACTG-3', R-5'-ATGCCAGTGAGCTTCCCGTTTCAAG-3'; Human *MEF2C* F-5'-TCCACCAGGCGCAAGAATACG-3', R-5'-GGAGTTGCTACGGAAACCACTG-3'; Human *BDNF* IV F-5'-GCTGCCTTGATGTTACTTTG-3', R-5'-AAGGATGGTCATCACTCTTCTCA-3'.

**Fluorescence in situ hybridization.** FISH detection of transcripts was performed by RNAscope assay (Advanced Cell Diagnostics) per the manufacturer's instructions. Target probes were either custom-synthesized or purchased from the available probe catalogue. Manufacturer's standard single red chromogenic/fluorescent (with the Fast Red fluorescent label) or multiplex fluorescent protocols were used for cultured human cells and fresh frozen human and macaque sections. Coverslips or sections were mounted with DAPI Fluoromount-G (Southern Biotech) for the visualization of nuclei. For negative controls, the probe against the bacterial gene *dihydrodipicolinate reductase* (*dapB*) was used, and no signal was detected in any of the experiments.

**Immunostaining and antibodies.** Neuronal cultures grown on glass coverslips were fixed with a solution of 4% paraformaldehyde and 4% sucrose in 1 $\times$  PBS pH 7.4 for 8 min at 27 $^{\circ}\text{C}$ , blocked for 1 h at 4 $^{\circ}\text{C}$  with 0.1% (w/v) gelatin and 0.3% (v/v) Triton X-100 in 1 $\times$  PBS pH 7.4 (GDB), and incubated overnight at 4 $^{\circ}\text{C}$  with the following primary antibodies diluted in GDB: anti-GFP (rabbit, 1:500, Life Technologies, A21311), anti-MAP2 (chicken, 1:1,000, Lifespan Biosciences, LS-C61805) and anti-MAP2 (chicken, 1:1,000, Abcam AB5392), anti-GFAP (rabbit, 1:500, Dako, Z033429-2), anti-SATB2 (mouse, 1:500, Abcam AB51502), anti-CTIP2 (rat, 1:300, Abcam, AB18465), anti-TBR1 (rabbit, 1:300, Abcam, AB31940). The *OSTN* antibody (rabbit, 1:500) was raised against a C-terminal region of *OSTN* (NP\_937827.1, amino acids 112–127, PKRRFGIPMDRIGRNR), then affinity-purified. All secondary antibodies were AlexaFluor-conjugated (Life Technologies). Coverslips were mounted with DAPI Fluoromount G (SouthernBiotech).

**Imaging and image quantification.** *In situ* and immunofluorescence experiments were imaged on either an AxioVision Imager Z1 or an LSM 5 Pascal (Zeiss). The individual GFP-positive neurons used for Sholl analyses were selected and imaged using an LSM 5 Pascal with a 40 $\times$  objective in a blinded manner. The neurons were traced using an ImageJ (NIH) plugin NeuronJ<sup>41</sup>, and Sholl analysis was performed by a blinded investigator using Sholl tool of Fiji<sup>42</sup>, quantifying the number of dendritic intersections at 10- $\mu\text{m}$  intervals from the cell body.

**hFBC transfections, chronic activity paradigm (CAP) and siRNA treatments.** The  $\sim 2\text{kbhOSTN:GFP}$  construct was transfected into DIV15 hFBCs using Lipofectamine 2000 reagent (Invitrogen) per the manufacturer's instructions. For RNA knockdown, we used a pool of four synthetic chemically modified ACCELL siRNAs (GE Dharmacon) to target *OSTN* and *MEF2C*. The pooled siRNAs, including a control pool consisting of scrambled sequences, were added to the medium for the last 3 days of culture (DIV27–DIV30, final concentration 1  $\mu\text{M}$ ) along with 20 mM KCl CAP treatment. The following siRNAs were used in the experiments: scrambled siRNA pool, UGGUUUACAUGUCGACUAA, UGGUUUACAUGUUUCUGA, UGGUUUACAUGUUUUCCUA, UGGUUUA CAUGUUGUGUGA; *OSTN* siRNA#1 pool, CUGUAGAUCACAAAGGUAA, GCUUCUUGAUGAUGAAUUGUG, GGAUCGGAUUGUAAGAAAC, CCUUUGAUUCUGGAGUCAU; *OSTN* siRNA#2 pool, ACAGCAAUAUGGAAGA, GCCUUCUGUAUGGAUU, UCUUUGGCUUCAUUG, CUCAGGAGUUGAAAGA; *MEF2C* siRNA pool, GGAUUAUGGAUGAACGUAA, CUCUUGUCUAAU AUUCGUC, GCACUAGCACUCAUUUAUC, CUGCCUUGUACUAAUGUUU.

**DNA constructs.** The 2-kb *cis*-regulatory regions directly upstream of the human *OSTN* (hg19 chr3:190,914,914–190,917,032) and mouse *Ostn* (mm10 chr16:27,305,609–27,307,640) transcription start sites were PCR amplified from genomic DNA using primers that incorporated SacI (5') and XhoI (3') nuclease recognition sequences. Luciferase constructs were cloned using the Firefly luciferase reporter plasmid pGL4.11 (Promega), which was linearized using SacI–XhoI digestion and then ligated with the mouse or human amplified DNA sequences. Further mutagenesis of pGL4.11–2kbhOSTN was performed using gBLOCKS (IDT) and Gibson Assembly Master Mix (New England Biolabs). The  $\sim 2\text{kbhOSTN:GFP}$  reporter construct was generated by cloning  $\sim 2\text{kbhOSTN}$



into pGL4.10[GFP] using SacI and XhoI restriction sites. pCAG-GFP (Addgene #11150) was used as a backbone for cloning the *OSTN* overexpression construct. pCAG-GFP was digested with the restriction enzymes EcoRI and PstI, while the PCR amplicon corresponding to the 'full-length' *OSTN* cDNA was cloned with EcoRI and SbfI. The predominant *OSTN* transcriptional variant sequence was determined by RNA-seq and 3'RACE analyses from depolarized hFBCs. The full-length *OSTN* cDNA is 3,219 bp long and consists of 5 exons. The first exon corresponds to the 5'UTR, the second exon starts with the initiation codon ATG (bold) and the last exon corresponds to the 3'UTR. The full-length *OSTN* cDNA sequence is as follows (alternating exons are underlined): AGGGCTGAGTTTTGGAGAACTGCAGAGACAGTACTCTAAAGTTA GAATCTCTGATCTTTCACGAGATGCTGGACTGGAGATGGCAAGTG CACATTTCTCCTGGTCTGACACTGACACTGTGGAGCTCAGGAAA AGTCCTCTCAGTAGATGTAACAACAACAGAGGCCCTTTGATTCTGGAGT CATAGATGTGCAGTCAACACCCACAGTCAGGGAAGAGAAATCAGCCA CTGACCTGCAGCAAACTCTTGCTTCTTGATGAATTGGTGTCCTAG AAAATGATGTGATTGAGACAAAGAAGAAAGGAGTTTCTCTGGTTTTG GGTCTCCCTTGACAGACTCTCAGCTGGCTCTGTAGATCACAAGGTA AACAGAGGAAAGTAGTAGATCATCCAAAAGGCGATTTGGTATCCCC ATGGATCGGATTGGTAGAAACCGGCTTTCAAATTCAGAGGCTAAATG CAACTTCCTTGGGTGAAATGTCACAGCAATATGGAAGATGCTTCACTG AAGTTATTCACATCTCTAATGATTAATCTTAAAGAACTGACCTTCT GCAATCTCTTCCAAATTTAGTTAACTTCACTGCTTCACTACATTCAGATTG TTACAGCTTCAATTAATTTGTGTAATCATTTTGATGCACGTACATTT TAA AATTATATATTTTAATTATTCAAGAATGGTTAACTTCCCTTAA CTTACTTTTAAAAATAATAATTAAATACACAATACAGTGAATGCCTTC TGTATGGATTACCATTGCACATGTTTGTAGTCAAAGAATAATAACAAA AGACAGATTTGCTTCTGTGAAATTTAGTTATAAGTGTCCATTTATGGG AATGAGGAAAGGCAATGCTGTGATTTTCTGTTGAGTACTTTCACCTCC CTGTATTCATTTTCCAAGAGTCTGATCGGTAATAAT TATGAAATTA GGCTTCTCTTTTCATATTCAGTTTTCAGTCATGTTTCAGAAAAATAAAA CACAGCCCCAATGAGCTATTGACTAGAAATTAAGAAAGTGAAGGACA TTACTACTTTGTCAAAGTTAGATACATGTTCCCTTAAAGATTTCCAT TTTTCTAAATCTGACAGTTAAGAGCAGTAGTGTCTCATTAGGAGGGGA GTAAGCTCACACAGAGGTAAATGAAAGTAGGAGGGAAGTCAAAGAA TTACCACCAAGACAGGTTAGGACCAGCTAAGCACACATCATTTTAGCT CAGTACACTTCAGCATACATGATGATCTTTTGTATATCTTGGATTA ATCTAAGAACTGTTTACTGTGTTTTCATAGTATGGCTTCTTGGCTTCA ATTGTCTTATTATCCTTAGTAAGCCAATTGAAGAGCATAATGATTTTGAG AATGATTTCTTAAATATCATTAGATTTTGAATGACTTATTAATTAAC ATAAGTTTTCGTATTGTAGAAACTCAGTTCTCAGTAATAACTATGATG TTACTGTAGCTTGGACACATAGGTCATTTGTCATTTGGATATACTTTGA AAACACACAAAAAACTTTCTATGGAACAGAGATTCATCATAAGTTA CTTAGCAGAAGTTTATAAAGCATCGAAAAACACTTC CTCTGTAAACCC TAAAAATCACTGCTGATACGTGGGAGGAAAAAGTTTTGTCCAGTAG AGCAAAGGCTTATTTTCAGCATAAAAAAGAGAGTGTGTGAGTTGTGAGA AGGTGTCTTAATTTAAAGGAAGAGGAGAGAAATGGGCAATTTGCTAAT CTTAATCAAAAAATTAAGGACTTGTATGATCCAAAGATAAAAAAATCA ACTCAAAGAAACATAACAAAAATATATTAGAAAAAGAGTTAACTAAGAAA TTGACCTTTATGAAGGACTAATTTTTCAGTTATCCAATGTGCTTTTAA AGATAAAGTTTAAAAAGAGGCAAAAAAGAGGTGAAAGAAATGAACATGTG TACTTAGAATATATATATTTGGTTTGTATGTAATCTTCTAGTTTCTTGG ATATTTTAAAAATGAAAGTCTTCAAAATTAATTTAGGAAAAATAGTAA ATAATTTTTTTTCATTTGTAGTAGAGGGCTGTAGCCAAGAGAAGTGAATT TTGGGAAATCTAACCAATCTTTTTTTCATAAACTTGTGTAGAAAGCT AGCAATAGCACATAGGGAAGTATCCTGGAGGACTCAGGAGTTGAAAGA ATTTATTAAGAAGTTGTGATATCTTGTCTGCCTTTCCCCAACCTTTGTG ATGAACATAAGAAATGTTTCTTCTATTTAGAGCTCTTCTTCTCTTG TATGCACATTTGGGTATTTGTAAGCTTCTAATCCAATTTGGGTTCTGCTTC ATGCTTTGACAAAAGGTATTAACCTACTTTAGCCTAAACTTTCTCAT GGTAATATTTGGAAGTTGGATTATGCAAAATGATTTCTCCATCTTTTA TTTTTTAATTCAAAATTAGACTATGACATCCAATTAATTAATAAAGAA GTCACATATGATGATTATGCTATAGTTTCAATGTGTAATGATTTTAC CTAATATACACACAATTTCCATGAAGGGAAGAAAATGTTTTCTCCACTT ATAATCTATTTTATTTTCAATTTTAAATTTTACCCTACTTCAATTCAGA GTAGAAAATAAGTCAGCAATATACTAAATAATGGGGCTATTCTTTTAAC ATAGCAGTAATTAAGACAGAAATTTTGTAAAGAAATGACAAGTCATC TCACCTTTATTTCAATGCAATAGGTAATTAATTAATTAATTAATTTCAA CTTGAAGGGACTTTTTTGTGTTTCAAAATAATGCATTACTTTTTTCT TCTTTGCTTCTGTATGAACCTTTATAGAGCAATGAATATATGTATATGG AGTTCTGGGTTCTAGTGTCAATTACATAATCAAAATTTCAATAAAGGAT GTTAGTTACTGGCTATGTTGCTCTAAATTTACACACTAAAAAATGT

CTGTCAAGTTGTACCTTTAACTGTTCATAGCTTTAGGGAATTAAGTTT CTTAAACCAAAATTATGAAAAATAACTTAATGGAATCTTCTAAAAGGA AAAAGTATAAAAAGCTTTCTGAATGATATTACCCCTTATACCTAAAGGC TCAAGATGCTTGAATATG GTTCAACTTTTCCAAAGTTAATAACAAGGG ATGATGAAA. Further construct sequences are available upon request.

**hiPSC differentiation to excitatory cortical neurons.** Previously characterized healthy control hiPSC lines 20b and 18a (ref. 43) were maintained in mTEsR medium (StemCell Technologies) on hESC-qualified matrigel (Corning)-coated tissue culture plates and passaged using Dispase (1 mg/ml, Life Technologies). Cell lines were mycoplasma-negative by PCR (LookOut Mycoplasma PCR Detection Kit, Sigma). hiPSCs were differentiated into dorsal telencephalic neural progenitors using a previously published protocol<sup>44</sup> without inducing sonic hedgehog signalling (no SHH, see Extended Data Fig. 5). After 18 days, cultures were enzymatically dissociated to single cells using Accutase (StemPro Accutase, Life Technologies) and were replated on Growth Factor Reduced matrigel (1:30, Corning) at 10,000 cells/cm<sup>2</sup> in human neurobasal (hNB) medium supplemented with 10  $\mu$ M ROCK inhibitor (Y27632, Tocris). hNB media was replaced 24 h later and supplemented every 2–3 days thereafter. Dissociation and replating (100,000 cells/cm<sup>2</sup>) was repeated at day 40. Cultures were silenced on day 81 and then KCl-depolarized (see above) and harvested on day 82. hNB medium: neurobasal medium (no glutamine), with 1  $\times$  penicillin/streptomycin, 1  $\times$  GlutaMax, 1  $\times$  MEM non-essential amino acids, 1  $\times$  B27-supplement without vitamin A (all Life Technologies), 1  $\times$  N2-B supplement (StemCell Technologies), 1  $\mu$ M ascorbic acid (Sigma), 20 ng/ml rhBDNF and 10 ng/ml rhGDNF (PeproTech).

**Luciferase assays.** Luciferase assays were performed using the Dual-Luciferase Reporter Assay System (Promega), and all constructs generated in this study were cloned using pGL4.11[Fluc] (firefly luciferase reporter construct – *Fluc*) and co-transfected with pGL4.74[Ren] (Renilla luciferase expression construct) as an internal transfection control. E16.5 C57BL/6 mouse embryonic cortical cultures were dissected and cultured as previously described<sup>45</sup>. Briefly, E16 mouse cortices were dissected, dissociated using a 1:100 dilution of papain suspension (Worthington, LS003126), rinsed with a 0.6% (w/v) solution of Ovomucoid Trypsin Inhibitor (Worthington) and BSA (Sigma) in HBSS and triturated briefly to dissociate cells, and single cells were plated at 150,000 cells/cm<sup>2</sup> on poly-L-ornithine- (Sigma) and laminin (Life Technologies)-coated tissue culture plates in mouse neurobasal (mNB) medium. mNB medium: neurobasal medium (no glutamine), with 1  $\times$  penicillin/streptomycin, 1  $\times$  GlutaMax, 1  $\times$  B27 supplement (all Life Technologies). At DIV5, control and experimental plasmids were transfected into the cultured cells using Lipofectamine 2000 (Life Technologies) following the manufacturer's protocols. On DIV6 cultures were silenced, and on DIV7 cultures were KCl-depolarized (see above) for 6 h before washing once with cold 1  $\times$  PBS and collecting cells in lysis buffer. For luciferase assays, hFBCs were transfected with plasmids on DIV25, silenced/treated on DIV27, and stimulated and harvested on DIV28. Protein lysates were analysed using Promega reagents according to Assay System instructions on a BioTek synergy 4 microplate reader, Gen5 1.11. Firefly luciferase activity readings were normalized for each experimental replicate using the Renilla luciferase activity reading from the same sample: Fluc/Ren. Two or three normalized Fluc values (experimental replicates) were averaged for each biological replicate value. The average normalized Fluc value of each condition treated with KCl was divided by the average normalized Fluc value of that same condition left untreated to obtain the fold-induction of the Fluc reporter for each condition in each biological replicate: avg. +KCl / avg. –KCl. Statistics were performed using a one-way ANOVA and Holm–Sidak test for multiple pair-wise comparisons. For CsA/FK506 treatment comparisons to DMSO treated controls, a one-way unpaired Student's *t*-test was applied. All constructs were tested from at least two independent plasmid preparations. Control plasmids used: '3xMRE' (generously contributed by the Eric Olson Laboratory) and '3xMREmut', mutated to inactivate the MRE sites as previously described<sup>18</sup>.

**Drug treatments.** Calcineurin activity was inhibited by addition of DMSO solutions of cyclosporin A (10 mM) and FK506 monohydrate (1 mM) (both Sigma) at 10,000  $\times$  dilutions in NB medium, for final concentrations of 1  $\mu$ M and 0.1  $\mu$ M in the culture medium, respectively. DMSO was used as the vehicle control. Cultures were treated at the same time as silencing treatment and were incubated for 18–20 h before KCl depolarization (see above).

**Monocular inactivation assay.** All procedures conformed to USDA and NIH guidelines and were approved by the Harvard Medical School Institutional Animal Care and Use Committee. Owing to the robust nature of this assay and scarcity of these specimens, we performed this experiment on two animals. One eye of each of the two adult male macaques (monkeys #1 and #2) was inactivated for 22 h by TTX injection as follows: the animal was anaesthetized with 10 mg/kg ketamine.

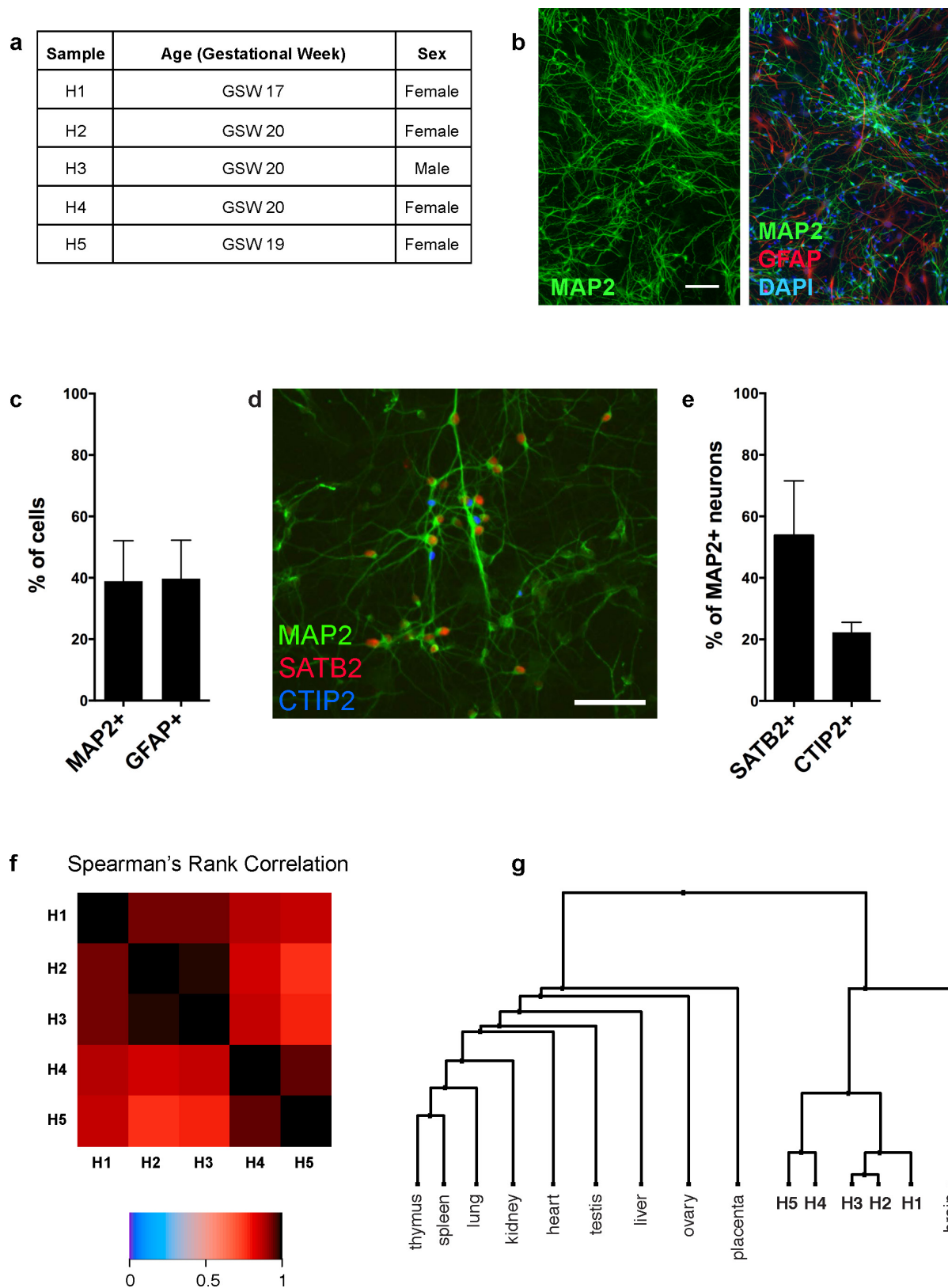
Then 10  $\mu$ l of sterile TTX solution (Tocris, 4.7 mM in sterile saline) was micro-injected at a rate of 1  $\mu$ l/min into the vitreous of the right eyeball using a sterilized Hamilton 0.5 inch 30 gauge 10  $\mu$ l microsyringe inserted through the sclera and controlled by an automated syringe pump. The microsyringe was kept in place for an additional 5 min to allow the pressure to equalize before retraction. The pupil of the injected eye dilated after the injection, indicating appropriate TTX delivery. The animal was immediately brought back to its home cage, allowed to recover from anaesthesia, and monitored post-operatively. On the following day, to maximize visualization of possible *OSTN* transcripts, 6 h after the beginning of the facility's normal light cycle (22 h after the TTX injection) the animal was anaesthetized with 15 mg/kg ketamine i.m. and then given a lethal dose of euthasol (pentobarbital + phenytoin) i.v. Immediately after death, brain tissue was removed and frozen on dry ice for cryosectioning into 15–25- $\mu$ m slices.

**Human brain sections.** The human fetal tissue used for *in situ* hybridization was acquired from the laboratory of N. Sestan at Yale University. Tissue was fixed using TissueTek VIP fixative for 24 h, then cryoprotected by immersing it in first 15%, then 20%, and finally 30% sucrose in RNase-free PBS at 4 °C. Samples were kept in each sucrose solution until completely equilibrated, when they sank to the bottom of the incubating vessel. Following cryoprotection, tissue was frozen using isopentane and dry ice. Frozen tissue was cut by cryostat into 12- $\mu$ m slices.

**ChIPseq of human neurons.** The ChIP assay in Fig. 2e and Extended Data Fig. 7d was performed as described<sup>17</sup> using hFBCs at DIV21 and mouse cortical cultures at

DIV7 under CAP conditions. H3K27Ac (Abcam ab4729), pan-MEF2 (Santa Cruz Biotechnology (C-21) sc-313) and MEF2C (Proteintech 18290-1-AP) antibodies were used.

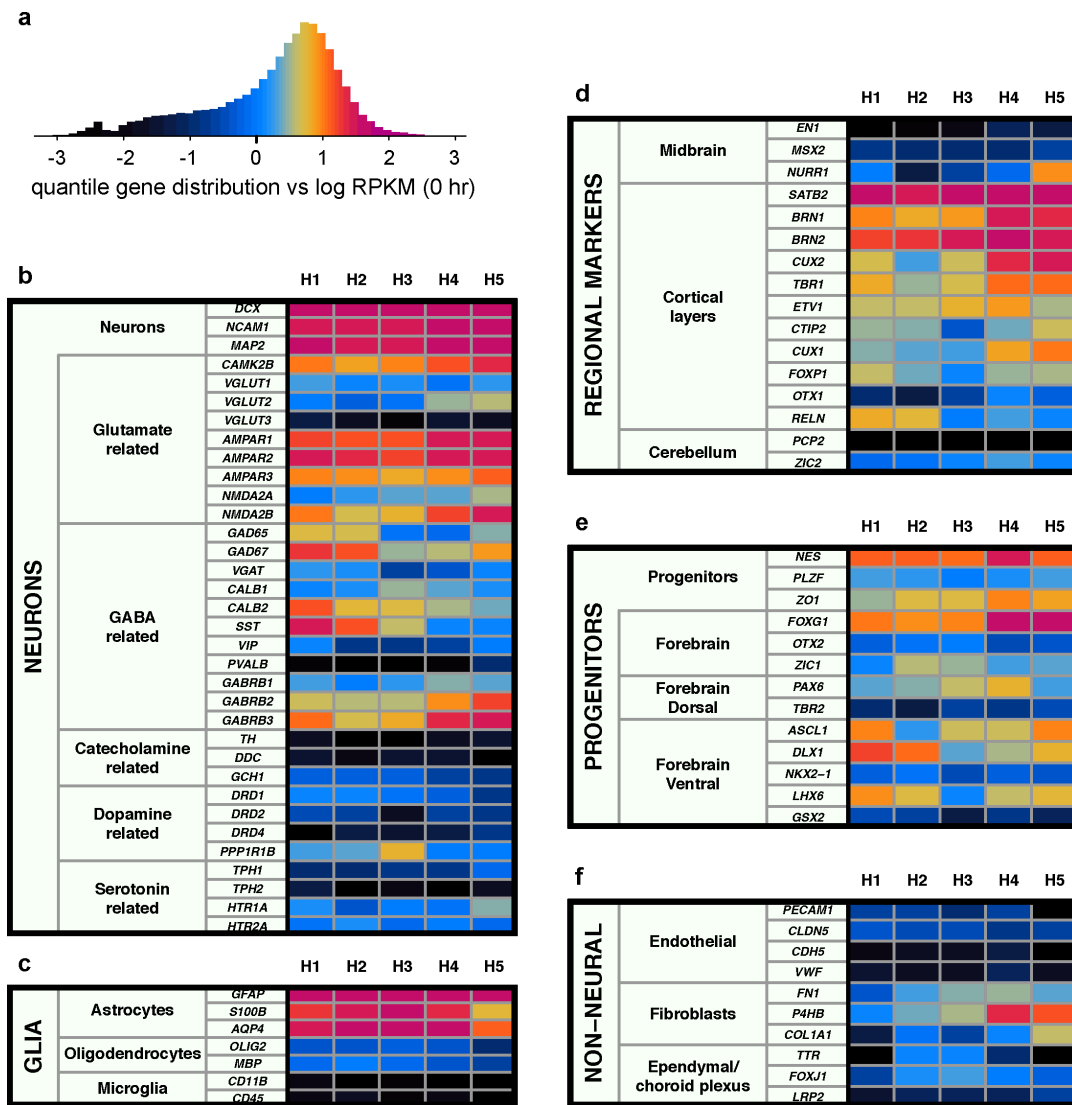
36. Brewer, G. J. & Torricelli, J. R. Isolation and culture of adult neurons and neurospheres. *Nat. Protocols* **2**, 1490–1498 (2007).
37. Gray, J. M. *et al.* SnapShot-Seq: a method for extracting genome-wide, *in vivo* mRNA dynamics from a single total RNA sample. *PLoS One* **9**, e89673 (2014).
38. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
39. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
40. Spiegel, I. *et al.* Npas4 regulates excitatory-inhibitory balance within neural circuits through cell-type-specific gene programs. *Cell* **157**, 1216–1229 (2014).
41. Meijering, E. *et al.* Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. *Cytometry A* **58**, 167–176 (2004).
42. Ferreira, T. A. *et al.* Neuronal morphometry directly from bitmap images. *Nat. Methods* **11**, 982–984 (2014).
43. Boulting, G. L. *et al.* A functionally characterized test set of human induced pluripotent stem cells. *Nat. Biotechnol.* **29**, 279–286 (2011).
44. Maroof, A. M. *et al.* Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. *Cell Stem Cell* **12**, 559–572 (2013).
45. Malik, A. N. *et al.* Genome-wide identification and characterization of functional neuronal activity-dependent enhancers. *Nat. Neurosci.* **17**, 1330–1339 (2014).



**Extended Data Figure 1 | hFBCs are mixed neuronal cultures that show high reproducibility.** **a**, Gestational week and sex of hFBC samples used in profiling of activity-dependent gene expression. **b**, Representative images of DIV6 hFBC neurons immunostained with the neuronal marker MAP2 alone or together with the glial marker GFAP and DAPI nuclear dye. Scale bar, 75  $\mu\text{m}$ . **c**, Quantification of MAP2- and GFAP-positive cells in hFBCs. Mean  $\pm$  s.d. from three independent cultures shown. **d**, Representative image of hFBC neurons immunostained with MAP2 (green), SATB2 (red) and CTIP2 (blue). Scale bar, 57  $\mu\text{m}$ . **e**, Quantification

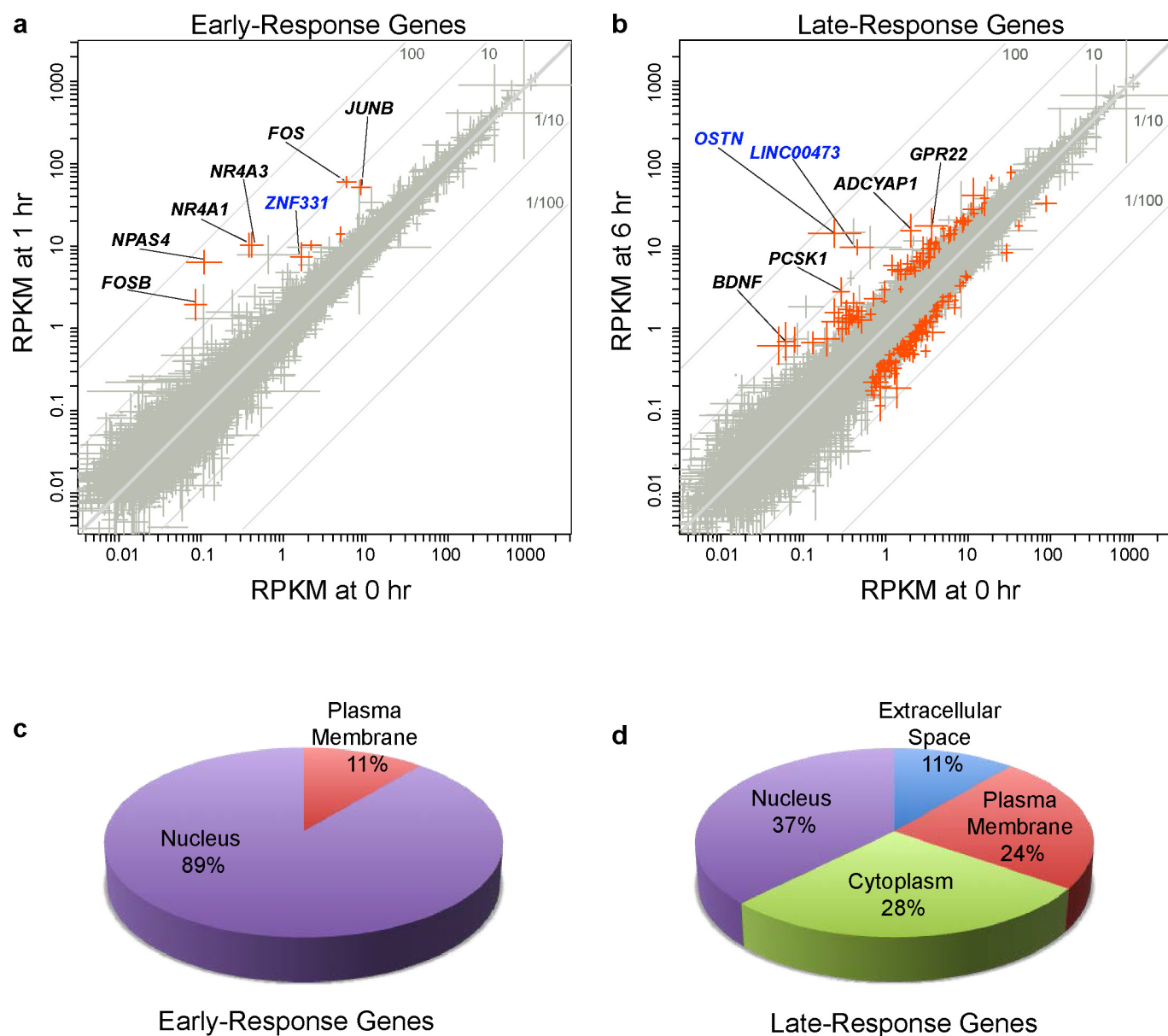
of the SATB2- and CTIP2-immunoreactive subpopulations of MAP2-positive hFBC neurons. Mean  $\pm$  s.d. from three independent cultures shown. **f**, Heatmap showing the Spearman correlation  $r_s$  of coding gene expression profiles among five biological hFBC replicates (H1–H5) (unstimulated neurons). **g**, Dendrogram of correlations among the gene expression profiles of hFBC replicates (H1–H5) and 10 human tissues, including whole brain<sup>37</sup>, based on hierarchical clustering with distance measure  $1 - r_s$ .





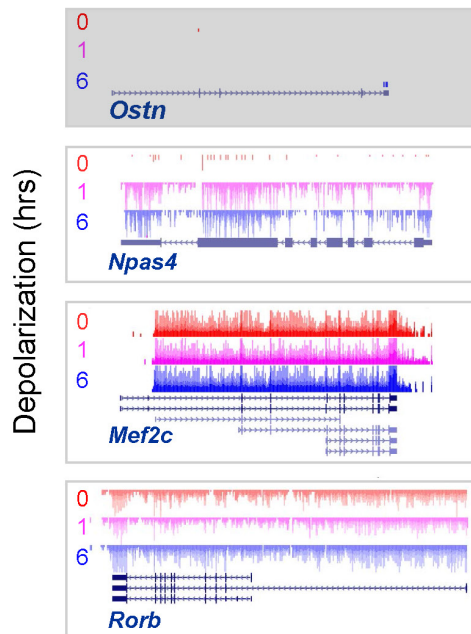
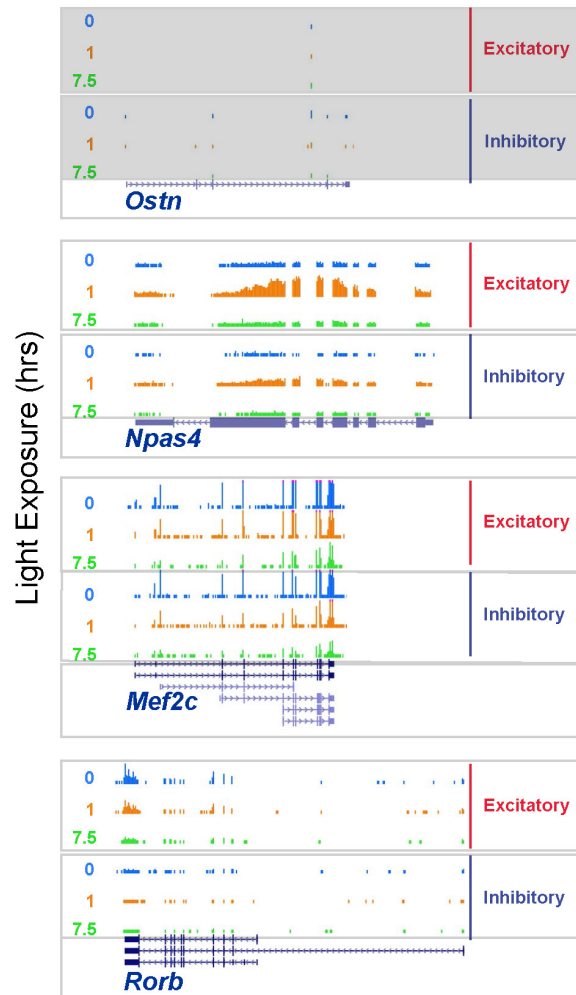
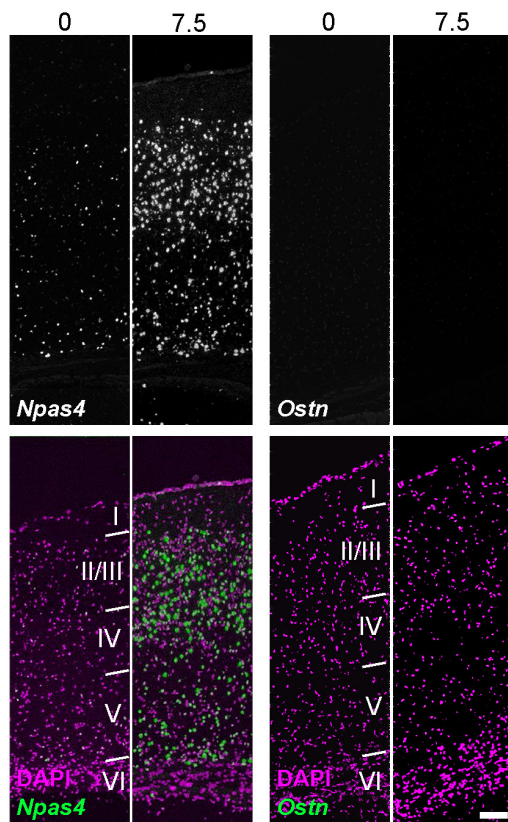
**Extended Data Figure 2 | hFBCs are mixed neuronal cultures with a substantial representation of cortical neuronal subtypes.** **a**, Quantile distribution constructed from combined log-gene expression levels of five unstimulated hFBC samples, with associated colour scale (see Methods).

**b–e**, Expression of selected marker genes classified by neuronal subtype (**b**), glial cell type (**c**), brain region (**d**), neural progenitor cell type (**e**), and non-neural cell type (**f**).



**Extended Data Figure 3 | RNA-seq profiling of activity-dependent gene expression in human neuronal cultures.** **a, b,** RNA-seq analysis of membrane depolarization-induced hFBC gene expression changes after 1 h (**a**) or 6 h (**b**). Scatterplots depict the geometric mean of genes' non-zero expression values  $\pm$  s.e.m. from five independent hFBC cultures. Fold change is proportional to distance from the diagonal. Genes passing filters for expression and significant activity-dependent changes are highlighted in red (BH-corrected  $P$  values controlled for FDR  $\leq 0.15$  based on a

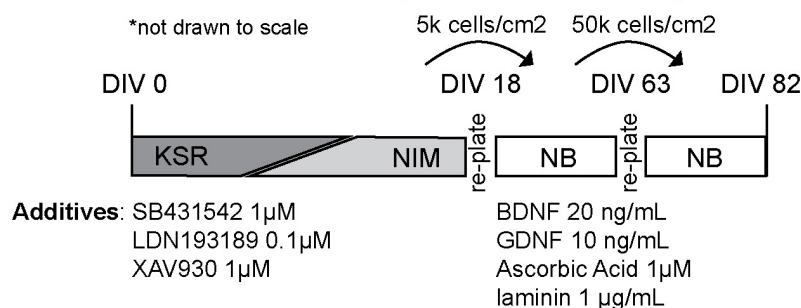
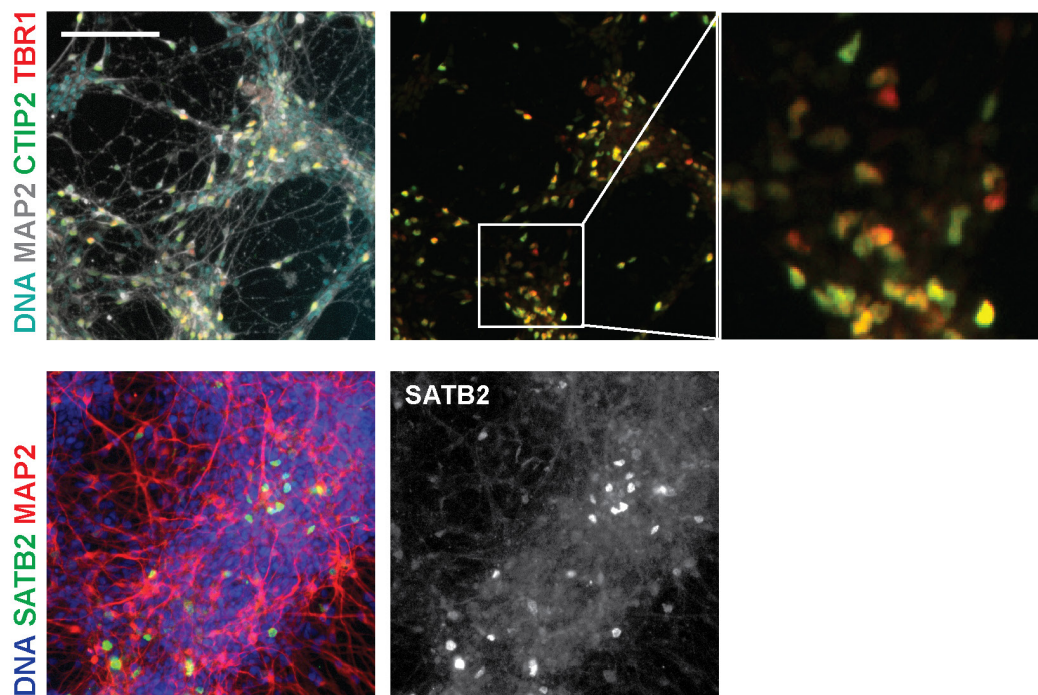
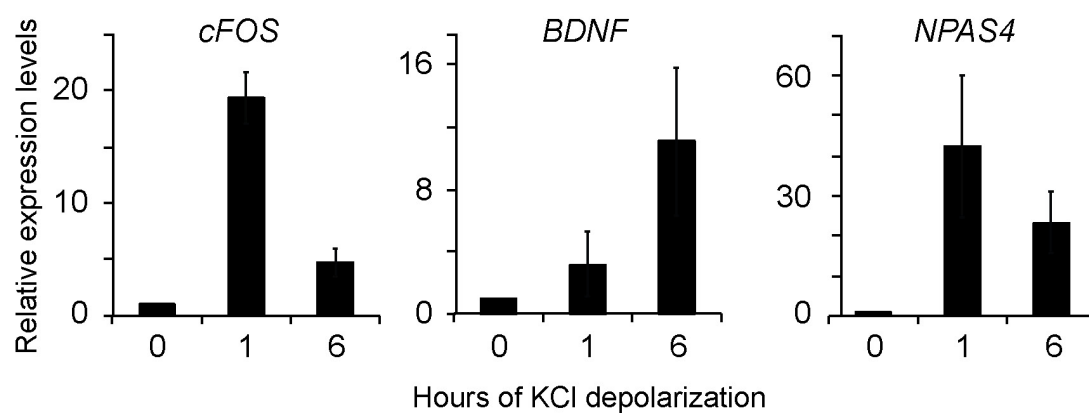
negative binomial model<sup>38</sup>, magnitude of change (ratio  $\geq 2.0$  or  $\leq 0.5$ ), and above-background expression (RPKM  $> 0.57$ ) on either axis, total reads  $\geq 3$  per time point). Selected genes exhibiting activity-regulated expression in human neurons but not in mouse neurons are indicated in blue. **c, d,** Pie charts showing the predicted subcellular localization of hFBC activity-responsive gene products induced following 1 h (**c**) or 6 h (**d**) KCl treatment. Analysis was performed using Ingenuity and GeneCards databases.

**a** Mouse Cortical Cultures (in vitro)**b** Mouse Visual Cortex (in vivo)**c** Light Exposure (hrs)

**Extended Data Figure 4 | *Ostn* is neither expressed nor activity-regulated in mouse cortical neurons *in vitro* and *in vivo*.** **a**, UCSC genome browser tracks for RNA-seq data from DIV7 cultured mouse cortical neurons depolarized for 0, 1 or 6 h with 55 mM KCl. The *Ostn* locus (grey) shows neither basal expression nor activity-dependent induction. The known activity-regulated gene *Npas4* shows clear activity-dependent induction at 1 and 6 h. Finally, the cortex-enriched transcription factor *Mef2c* and the layer IV marker *Rorb* show no significant expression changes in response to depolarization. **b**, UCSC genome browser tracks for RNA-seq data from visual cortices of dark-

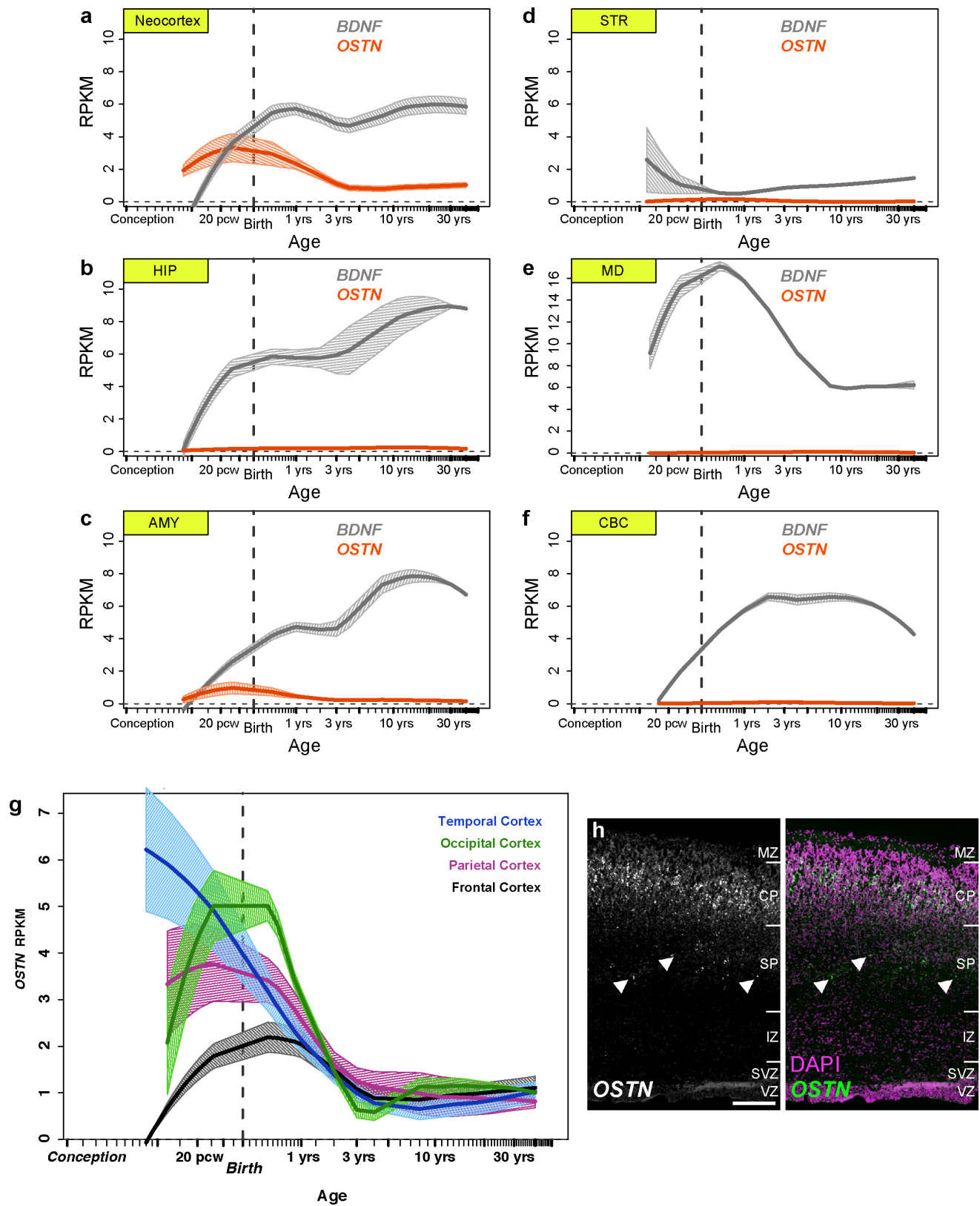
adapted (P42–P56) mice that were exposed to light for 0, 1, and 7.5 h. RNAs from excitatory and inhibitory neurons were isolated through the expression of a RiboTag transgene using *Emx* and *Gad2* Cre-lines, respectively<sup>34</sup>. The *Ostn* (grey), *Npas4*, *Mef2c* and *Rorb* loci show similar responses as in **a**. All genome browser tracks y-axis min = 0 and max = 10. **c**, FISH images of radial sections from primary visual cortex of dark-adapted (P42–P56) mice exposed to light for 0 and 7.5 h. Upper panels, grey-scale images of *Npas4* (left) and *Ostn* (right) probes. Lower panels, green-coloured images from upper panels, with nuclei marked with DAPI (magenta). Scale bar, 110 μm; cortical layers I–VI are indicated.



**a** Human iPSC-derived cortical excitatory neuron differentiation protocol**b****c**

**Extended Data Figure 5 | Differentiation and characterization of human iPSC-derived cortical neurons.** **a**, Schematic of the iPSC cortical neuron differentiation protocol<sup>44</sup> (see Methods). **b**, Immunostaining of DIV82 iPSC-derived neurons shows expression of cortical layer markers TBR1 (layer VI), CTIP2 (layer V), and SATB2 (layers II–IV).

**c**, Quantitative RT–PCR analysis of known activity-dependent genes from DIV82 iPSC-derived neurons 0, 1, and 6 h after membrane depolarization with 55 mM KCl. Data shown as mean ± s.e.m. from two independent iPSC lines. Scale bar, 100 µm.

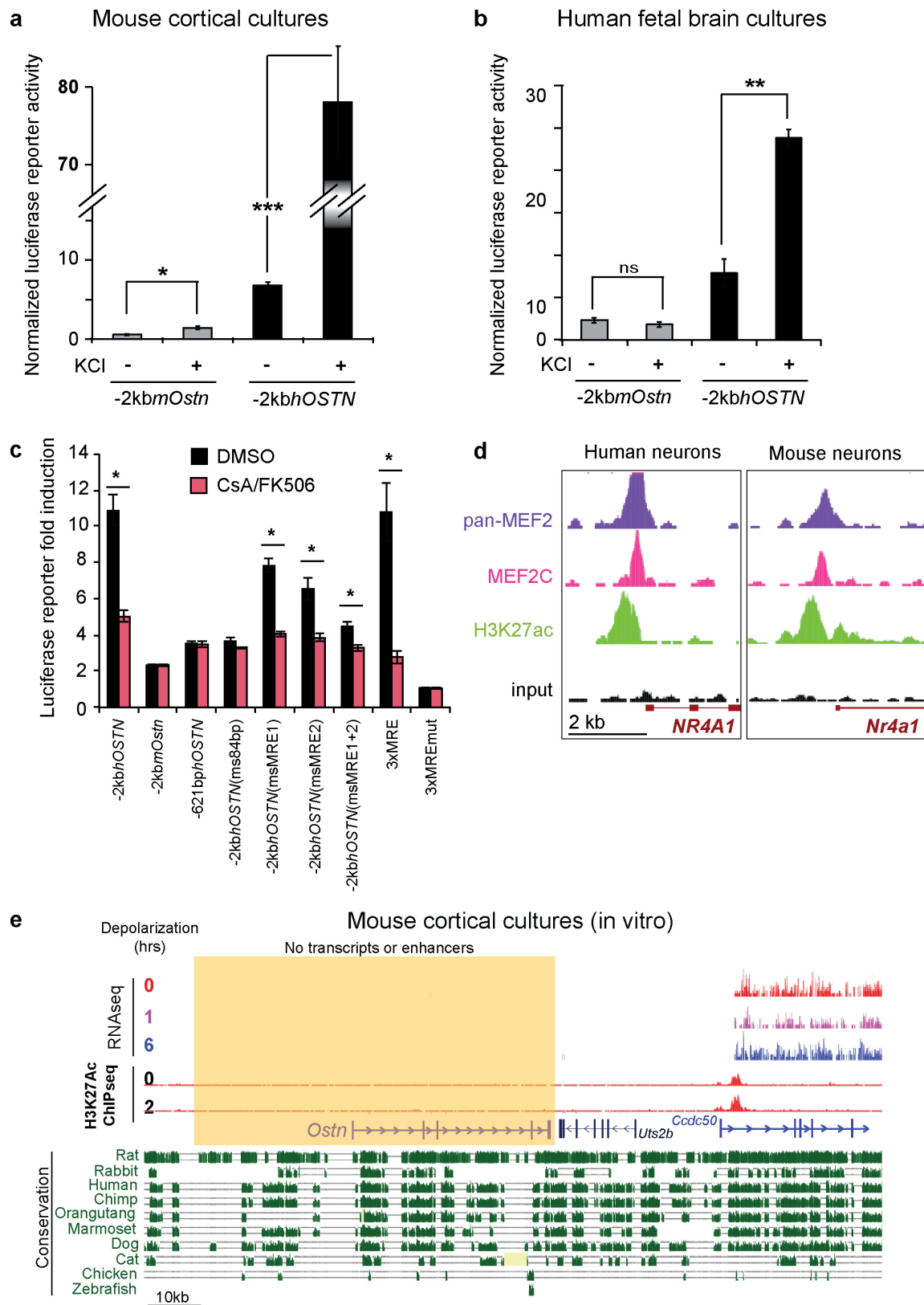


Extended Data Figure 6 | See next page for caption.

**Extended Data Figure 6 | *OSTN* is primarily expressed in the neocortex of human brain.** BrainSpan (<http://www.brainspan.org>) RNA-seq data showing expression levels of *OSTN* (red) and *BDNF* (grey) in 6 human brain regions (a–f; neocortex, hippocampus (HIP), amygdala (AMY), striatum (STR), mediodorsal nucleus of the thalamus (MD), and cerebellar cortex (CBC)) and *OSTN* in subregions of the human cortex from 8 pcw through 40 years old (g). Loess-fit curves depict mean

expression with bands showing one s.e.m. **h**, FISH images showing *OSTN* expression in a radial section of human fetal brain (pcw16) illustrating selective enrichment of *OSTN* in the developing cortical plate of the paracentral lobule. Isolated *OSTN* signal also appears to be localized to migrating neurons (arrowheads) of the subplate. Scale bar, 200  $\mu$ m. MZ, marginal zone; CP, cortical plate; SP, subplate; IZ, intermediate zone; SVZ, subventricular zone; VZ, ventricular zone.

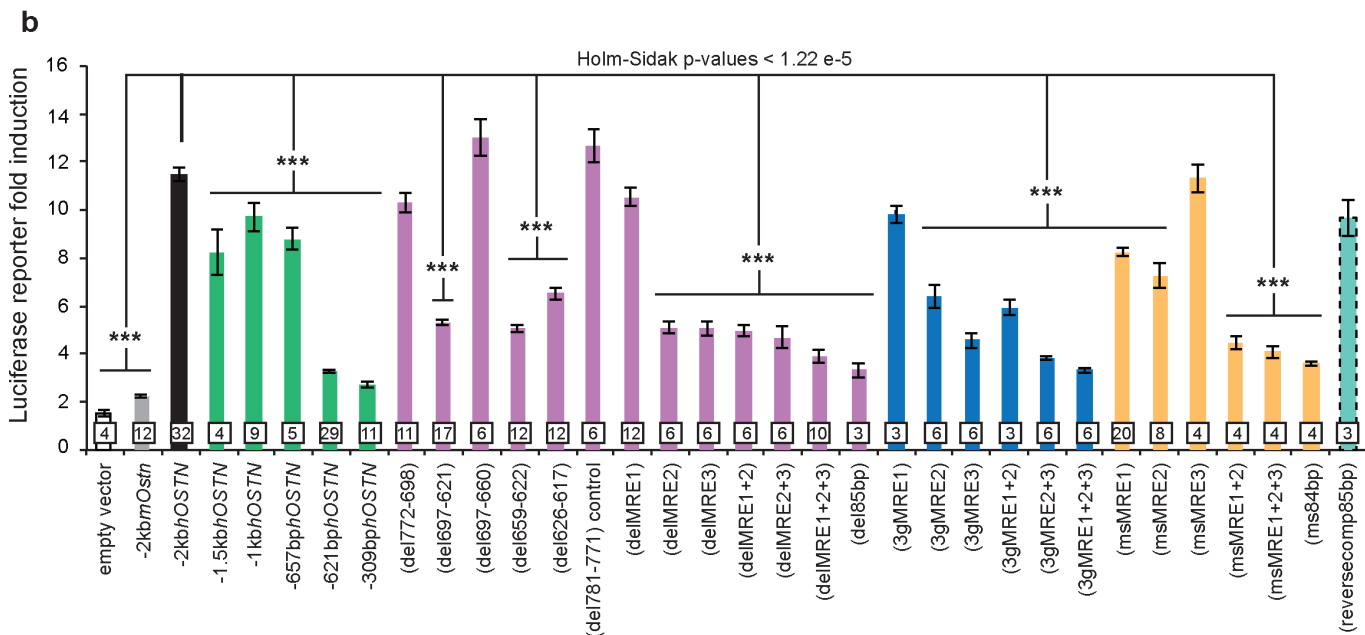
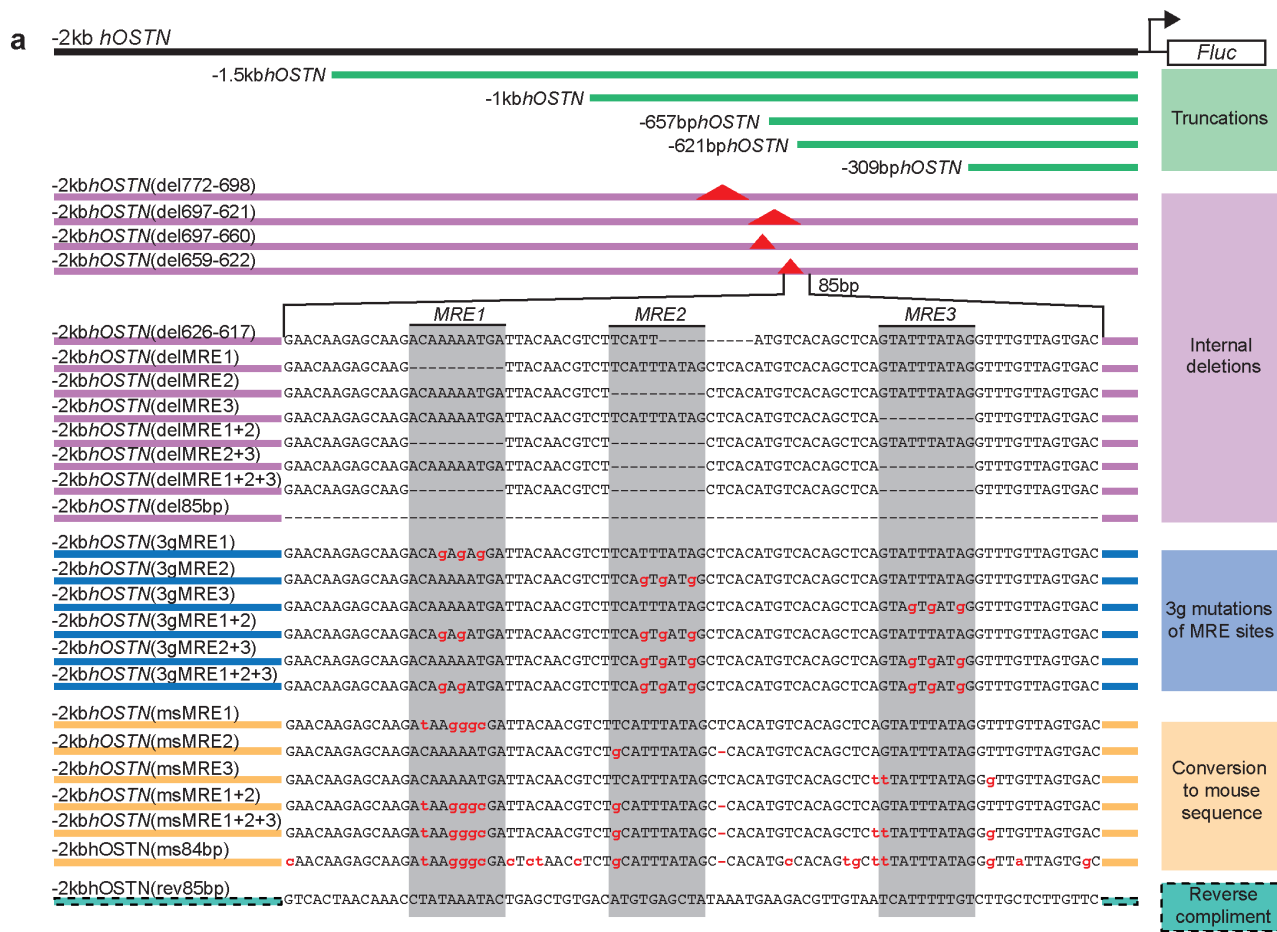




Extended Data Figure 7 | See next page for caption.

**Extended Data Figure 7 | Luciferase and ChIP assays in human and mouse neurons.** Direct comparison of the ability of the human and mouse –2kb regulatory sequences to drive reporter expression in mouse (**a**;  $n = 8$ ) and human (**b**;  $n = 3$ ) neuronal cultures in response to KCl depolarization.  $n$  = number of biological replicates. Mean normalized firefly luciferase activity (Fluc/Ren)  $\pm$  s.e.m., Student's  $t$ -test, NS, not significant,  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ . **c**, Luciferase assays performed in mouse neurons in the presence of calcineurin inhibitors (CsA and FK506, red) or vehicle (DMSO, black), Student's  $t$ -test,  $*P < 0.05$ . **d**, ChIP-seq using a pan-MEF2 antibody (purple), an MEF2C-specific antibody (fuschia), and an antibody specific for H3K27ac (green) in hFBCs (left) and mouse cortical neuron cultures (right) shows enrichment for MEF2 binding at the known MEF2-regulated gene *Nr4a1* (also known as *Nur77*). Y-axis scales here and in Fig. 2e were adjusted for

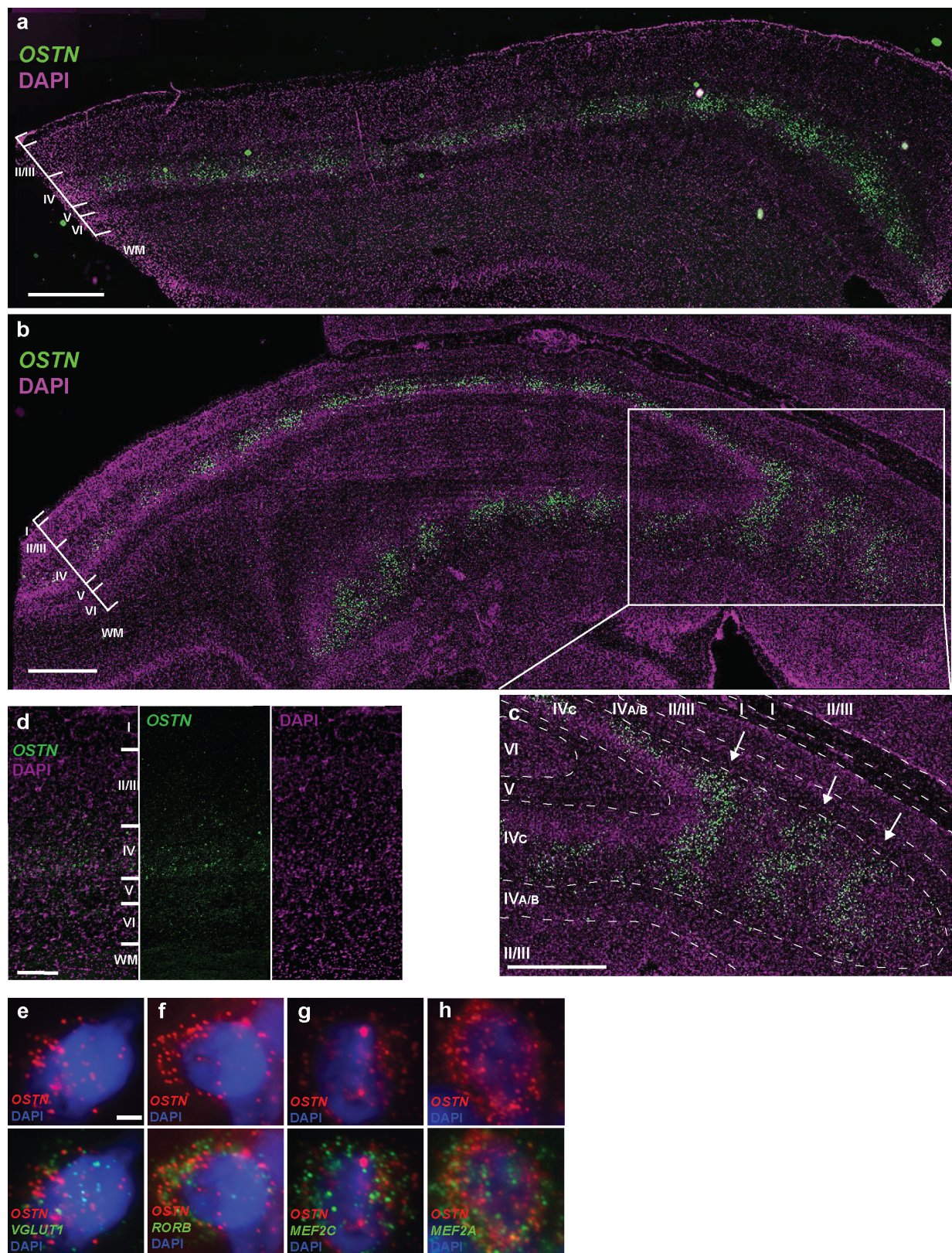
each experiment to normalize for variability in ChIP efficiency between these two different culture systems. We chose scales by setting MEF2 and H3K27ac enrichment to approximately equal levels at this positive-control locus, yielding all human tracks at max 10, mouse pan-Mef2, Mef2c, and input tracks at max 20, and mouse active chromatin (H3K27ac) at max 50. **e**, UCSC genome browser tracks for RNA-seq, ChIP-seq and vertebrate evolutionary conservation at the mouse *Ostn* locus, shaded yellow. RNA-seq tracks show no *Ostn* expression or induction in DIV7 mouse cortical neuron cultures following KCl depolarization (0, 1 and 6 h). ChIP-seq tracks<sup>45</sup> show H3K27ac peaks that mark active *cis*-regulatory regions at two time points: 0 h and 2 h after KCl depolarization. The nearby genes *Uts2b* and *Ccdc50* are also shown for comparison. No active *cis*-regulatory sites were found surrounding the *Ostn* locus. H3K27ac tracks are shown with max 5 and RNA-seq tracks with max 10.



**Extended Data Figure 8 | Luciferase reporter constructs and the complete set of assays.** **a**, Detailed summary of all luciferase assay reporter construct variants of the human genomic sequence –2 kb upstream of the *OSTN* transcription start site. **b**, Summary of all luciferase assays performed in mouse cortical cultures. Categories of construct modifications are indicated and grouped by colour. Biological replicate

numbers are indicated on the graph. Significant differences tested for by one-way ANOVA  $DF = (33, 266)$  and  $P < 1.0 \times 10^{-13}$ . Pair-wise comparisons were made using Holm-Sidak test for multiple comparisons using an overall error rate of 0.05,  $***P < 1.22 \times 10^{-5}$ . All values are mean  $\pm$  s.e.m.



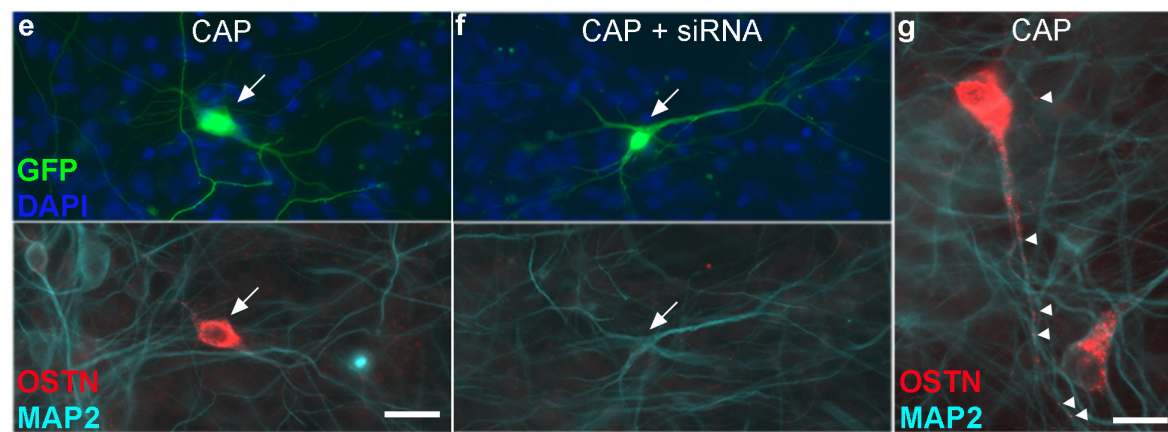
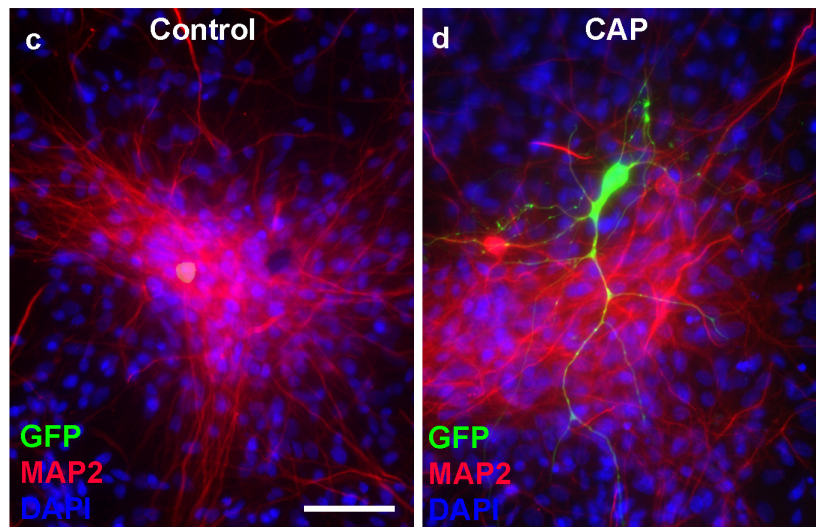
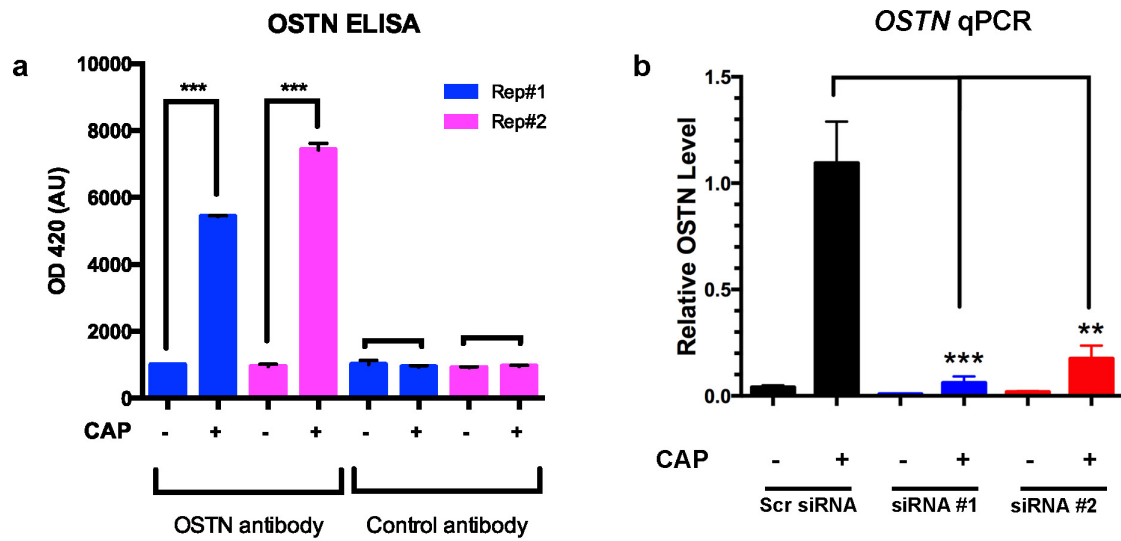


#### Extended Data Figure 9 | FISH for *OSTN* mRNA in macaque brain.

Layer IVC of active ocular dominance columns in primary visual cortex (V1) shows expression of *OSTN* after monocular inactivation of monkey#1 (a) and monkey#2 (b). Scale bar, 1,000  $\mu\text{m}$ . c, Expanded panel shows detail of partially tangential portion of tissue section in which *OSTN* is expressed in layer IVC ocular dominance column stripes. Scale bar, 1,000  $\mu\text{m}$ . d, *OSTN* expression is also enriched in layer IV of the multimodal parietal cortex. Scale bar, 250  $\mu\text{m}$ . e–h, Representative FISH images of layer IVC

neurons from the active columns, showing co-expression of *OSTN* with various cell-type markers, including *VGLUT1* (glutamatergic neurons, e;  $94.2 \pm 2.8\%$  of *OSTN*<sup>+</sup> cells were *VGLUT1*<sup>+</sup>,  $n = 170$ ), *RORB* (layer IV, f;  $93.3 \pm 6.1\%$  of *OSTN*<sup>+</sup> cells were *RORB*<sup>+</sup>,  $n = 92$ ), *MEF2C* (g;  $100\% \pm 0$  of *OSTN*<sup>+</sup> cells were *MEF2C*<sup>+</sup>,  $n = 148$ ), and *MEF2A* (h;  $100\% \pm 0$  of *OSTN*<sup>+</sup> cells were *MEF2A*<sup>+</sup>,  $n = 148$ ). Data are represented as mean  $\pm$  s.d. Nuclei are visualized with DAPI. Scale bar (e–h), 2  $\mu\text{m}$ .





Extended Data Figure 10 | See next page for caption.

**Extended Data Figure 10 | Biochemical detection and immunolocalization of endogenous OSTN protein in human neurons.**

**a**, ELISA quantification of secreted OSTN in the culture medium of hFBCs under CAP conditions in two biological replicates (Rep#1 and #2). Rat monoclonal anti-OSTN antibody and rat monoclonal anti-CD31 (control antibody) were used as the detection antibodies.  $n$  = number of biological replicates. Mean  $\pm$  s.e.m., Student's  $t$ -test \*\*\* $P$  < 0.001.

**b**, Quantitative RT-PCR analysis of *OSTN* induction in hFBC neurons treated with scrambled siRNA ( $n$  = 5) or two independent siRNAs against the *OSTN* transcript (#1;  $n$  = 5 and #2;  $n$  = 4) for three days in the presence and absence of CAP. *OSTN* expression is normalized to *GAPDH*.

**c, d**, Immunofluorescence images of DIV28 hFBC neurons transfected with  $-2kbhOSTN:GFP$  and left untreated (**c**) or maintained under CAP conditions (**d**). **e–g**, Immunofluorescence images of DIV28 hFBC neurons transfected with  $-2kbhOSTN:GFP$  (arrows) and treated with CAP for 3 days with (**f**) or without (**e**) treatment with siRNA targeting *OSTN*. Endogenous OSTN is predominantly localized in the soma and primary dendrites. Scale bar, 15  $\mu$ m. (**g**) Higher magnification of OSTN immunostaining after 3 day CAP treatment reveals punctate structures (arrowheads) in the dendrites. Scale bars, 48  $\mu$ m (**c, d**), 23  $\mu$ m (**e, f**), 15  $\mu$ m (**g**).



# Transplanted embryonic neurons integrate into adult neocortical circuits

Susanne Falkner<sup>1\*</sup>, Sofia Grade<sup>2,3\*</sup>, Leda Dimou<sup>2,3,4</sup>, Karl-Klaus Conzelmann<sup>5</sup>, Tobias Bonhoeffer<sup>1</sup>, Magdalena Götz<sup>2,3,4§</sup> & Mark Hübener<sup>1§</sup>

**The ability of the adult mammalian brain to compensate for neuronal loss caused by injury or disease is very limited. Transplantation aims to replace lost neurons, but the extent to which new neurons can integrate into existing circuits is unknown. Here, using chronic *in vivo* two-photon imaging, we show that embryonic neurons transplanted into the visual cortex of adult mice mature into bona fide pyramidal cells with selective pruning of basal dendrites, achieving adult-like densities of dendritic spines and axonal boutons within 4–8 weeks. Monosynaptic tracing experiments reveal that grafted neurons receive area-specific, afferent inputs matching those of pyramidal neurons in the normal visual cortex, including topographically organized geniculate–cortical connections. Furthermore, stimulus-selective responses refine over the course of many weeks and finally become indistinguishable from those of host neurons. Thus, grafted neurons can integrate with great specificity into neocortical circuits that normally never incorporate new neurons in the adult brain.**

Neuronal cell loss after brain injury or in neurodegenerative disease currently cannot be repaired. Promising results involve the transplantation of cells from various sources<sup>1,2</sup>, but it is not known whether grafted cells can truly replace lost neurons. Transplantations have been successful in improving clinical symptoms in Parkinson's disease, for example, but in this case fetal mesencephalic cells were transplanted into an ectopic location, the striatum<sup>3,4</sup>. Therefore, it is unclear whether there can be true functional replacement of lost neurons, requiring the proper integration of new neurons into pre-existing circuits at homotopic sites. Transplanted neurons have been shown to send out correct efferent projections and to form synapses with host neurons<sup>5,6</sup>, but little is known about the inputs onto transplanted neurons<sup>4,7</sup> and their participation in the processing tasks performed by the respective brain regions.

To address these questions, we chose layer 2/3 (L2/3) cells of the primary visual cortex (V1). They represent a main integration site for visual inputs, their connectivity is well known, and they have characteristic receptive field properties<sup>8</sup>. To ablate L2/3 neurons specifically we used a selective lesion method<sup>9</sup> and then transplanted late embryonic cortical neurons into the affected area.

## Grafted neurons have layer identity of lost neurons

We ablated L2/3 callosal neurons in V1 by retrograde labelling with chlorine e6 (Ce6)-coupled beads, followed by laser photoactivation-induced apoptotic cell death<sup>9</sup> (Fig. 1a, Extended Data Fig. 1a, b). Donor cells from mouse embryonic neocortex expressing fluorescent protein (Extended Data Fig. 2) were transplanted into this region 7–10 days later (Fig. 1a, b, Extended Data Fig. 3a, b). By four weeks post-transplantation (wpt), the overwhelming majority of new neurons had acquired morphologies typical of mature L2/3 pyramidal cells<sup>10</sup> (Fig. 1c, d, Extended Data Fig. 3c, and Supplementary Video 1). Most cells remained restricted to the transplantation site (Fig. 1b), and the majority (77%, 200 cells,  $n = 2$  mice) displayed L2/3 identity as demonstrated by Cux1 immunoreactivity (Extended Data Fig. 1c). To test for cell fusion

between transplanted and host neurons<sup>11,12</sup>, Emx1-Cre-driven donor cells expressing green fluorescent protein (GFP)<sup>13</sup> were grafted into a tdTomato reporter mouse line<sup>14</sup> (Extended Data Fig. 3d). The absence of GFP<sup>+</sup>/tdTomato<sup>+</sup> cells ( $n = 10$  mice, Extended Data Fig. 3e) and the elaborate morphological development of transplanted neurons (Supplementary Videos 1–3; see later) shows that cell fusion was not a confounding factor.

## Structural maturation of transplanted neurons

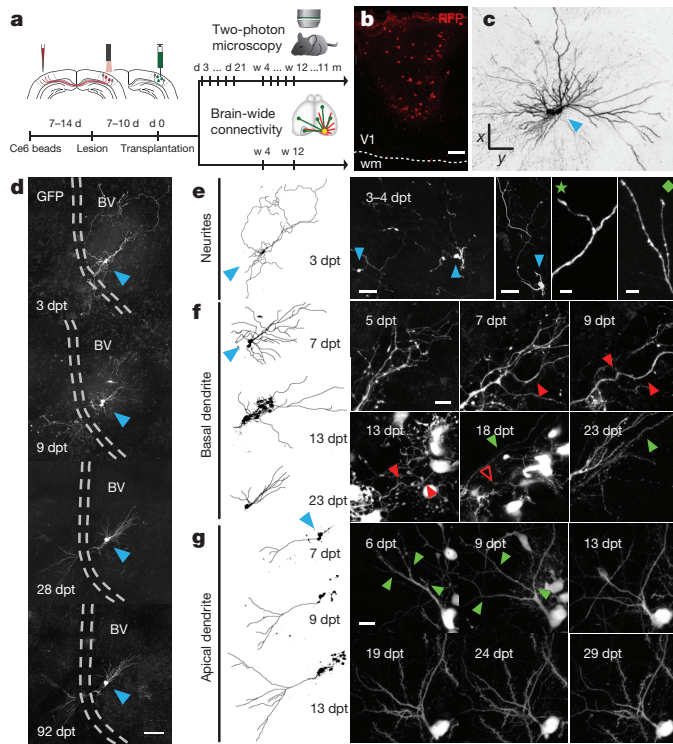
To follow the integration of transplanted neurons in V1 we used repeated two-photon imaging<sup>15</sup> (Extended Data Fig. 3a, b). Already by 4 wpt, many transplanted neurons displayed a fully mature pyramidal-cell-like morphology (Fig. 1c, d, Extended Data Fig. 3c, and Supplementary Video 1). Notably, neurons that had survived the initial phase and extended dendrites by 12 days post-transplantation (dpt) were present until the end of the imaging period (2 months: 97%,  $n = 6$  mice; 11 months: 94%,  $n = 2$  mice; Extended Data Fig. 4d).

Already by 3–4 dpt, neurons extended hundreds of micrometres of branched neurites, often with growth cones at their tips, some clearly identifiable as axons, based on the presence of first boutons (Fig. 1d, e, Extended Data Fig. 4b). At 5–6 dpt, a short, main apical dendrite featured primary and secondary branches, outlining the subsequent, more sophisticated structure of the apical dendrite (Fig. 1d, g). Thus, neuronal polarity is established by around 6 dpt ( $n = 6$  out of 6 mice < 7–9 dpt). Until the end of the second week after transplantation, apical dendrites grew to their full length and formed only a few additional higher order branches (Fig. 1g). By contrast, initial neurites later forming the basal dendrites underwent massive rearrangements until 4 wpt (Fig. 1f). Development of the basal dendrites included pruning as well as *de novo* growth of processes. Thus, similar to normal development<sup>10</sup>, the apical dendrite matures before its basal counterpart ( $n = 6$  out of 6 mice). Dendritic development is completed by 4 wpt, and the dendritic tree remained stable thereafter.

<sup>1</sup>Max Planck Institute of Neurobiology, D-82152 Martinsried, Germany. <sup>2</sup>Physiological Genomics, Biomedical Center, Ludwig-Maximilians University Munich, D-82152 Planegg, Germany. <sup>3</sup>Institute of Stem Cell Research, Helmholtz Center Munich, German Research Center for Environmental Health, D-85764 Neuherberg, Germany. <sup>4</sup>SYNERGY, Excellence Cluster of Systems Neurology, Biomedical Center, Ludwig-Maximilians University Munich, D-82152 Planegg, Germany. <sup>5</sup>Max von Pettenkofer Institute and Gene Center, Ludwig-Maximilians University Munich, D-81377 Munich, Germany.

\*These authors contributed equally to this work.

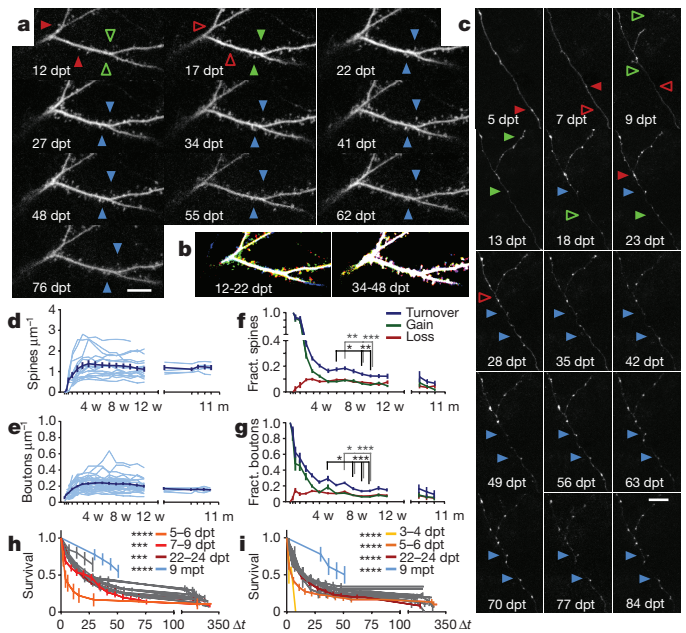
§These authors jointly supervised this work.



**Figure 1 | Transplanted embryonic neurons develop pyramidal neuron morphology.** **a**, Experimental procedures and timeline (see Methods). d, day; m, month; w, week. **b**, Transplanted neurons (RFP<sup>+</sup>) in host V1, at 5 wpt. wm, white matter. **c**, *In vivo* two-photon z-projection of transplanted neurons (GFP<sup>+</sup>) in host V1, at 45 dpt. **d**, Development of a transplanted neuron 3–92 dpt. Stable morphology between 28 and 92 dpt. BV, blood vessel. **e**, Left, inverted binary projection of neurites at 3 dpt of neuron in **d**. Right, transplanted neurons extending early neurites (blue arrowheads). Star, axonal growth cone; diamond, dendritic growth cone. **f**, Development of basal dendrites. Left, binary image; right, high magnification of neuron in **d**. Red filled and empty arrowheads, pruned process and former location, respectively; green arrowheads, newly formed processes. **g**, Development of apical dendrite. Left, binary of neuron in **d**; right, high magnification of another neuron. Early branch points remain stable (green arrowheads). **c–g**, *In vivo* two-photon z-projections. Blue arrowhead denotes location of cell body. Scale bars, 100  $\mu$ m (**b**), 50  $\mu$ m (**c–e**), 5  $\mu$ m (**e**, star, diamond) and 10  $\mu$ m (**f**, **g**).

### Formation of dendritic spines and axonal boutons

The elaborate morphology of dendrites and the long-term survival of transplanted neurons suggest their stable synaptic integration. As dendritic spines and axonal boutons are regarded as structural correlates of synapses, we followed the developmental dynamics of synaptic structures within days and up to 11 months post-transplantation (mpt; Figs 1a and 2, Supplementary Videos 2 and 3). Boutons were detected as early as 3 dpt (3–4 dpt; Extended Data Fig. 4b, c). By contrast, spines appeared first at 6 dpt but usually formed in the second wpt on already arborized dendrites (Extended Data Fig. 4a). Spine and bouton density increased greatly up to 4 wpt, and then subsequently plateaued (16 dendrites,  $n = 5$  mice:  $1.38 \pm 0.17 \mu\text{m}^{-1}$  (mean  $\pm$  s.e.m.)); 33 axons,  $n = 6$  mice:  $0.23 \pm 0.016 \mu\text{m}^{-1}$  at 4 wpt; Fig. 2d, e). Accordingly, the high initial turnover rates, largely based on a high gain of new structures and an increased, but less pronounced loss, decreased up to 4 wpt (Fig. 2f, g). Early-formed spines and boutons had a considerably higher probability of elimination compared to synaptic structures formed at 4–9 wpt (spines: 1.5–2.4 times; boutons: 2–8 times increased chance of elimination; Fig. 2h, i, Extended Data Fig. 4e, f). Turnover rates only stabilized after a subsequent period of transient increase between 5 and 8 wpt at a level of  $<15\%$  (spines:  $0.125 \pm 0.013$ ; boutons  $0.139 \pm 0.018$ ; both 10 wpt; Fig. 2f, g). By 8–10 wpt, neurons had developed largely stable pre- and postsynaptic structures. Spine densities and turnover rates at later stages ( $>8$  wpt) are comparable to reported data<sup>16</sup>.

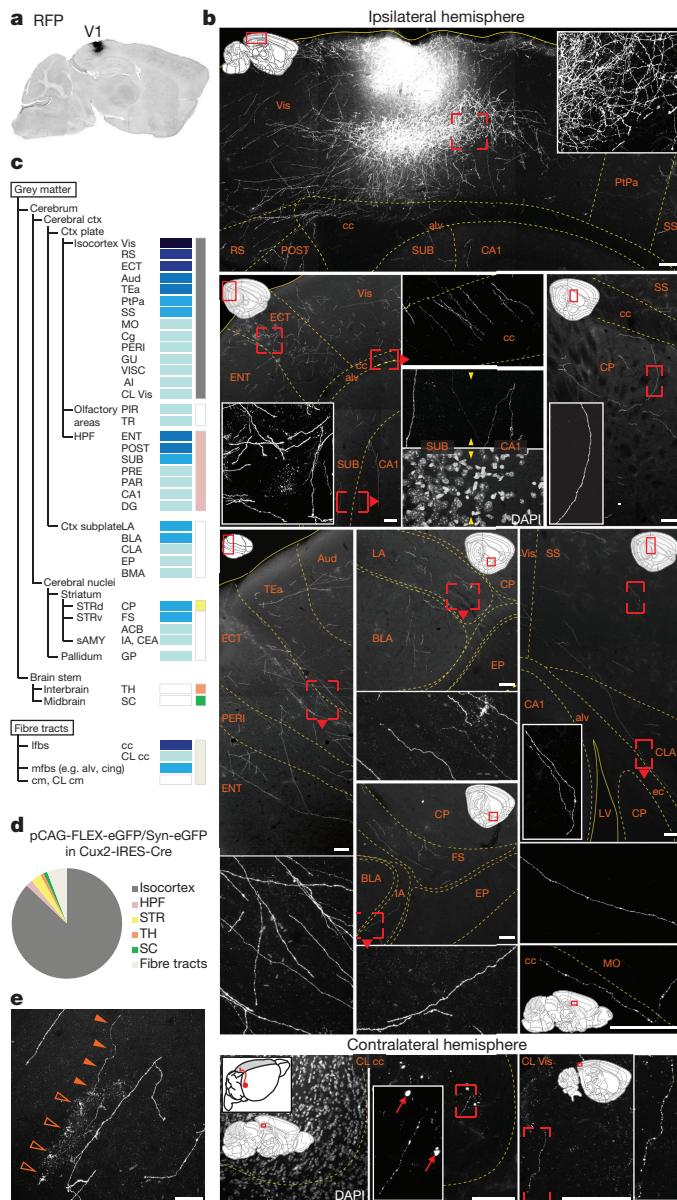


**Figure 2 | Transplanted neurons form synaptic structures.** **a**, **c**, Two-photon z-projections of dendrites (**a**) and axons (**c**). Arrowheads indicate lost (red), gained (green) and stable (blue) dendritic spines and axonal boutons. **b**, Early formed spines are dynamic, later formed spines are stable. Binary, colour-coded overlay of early (left; blue, 12 dpt; red, 17 dpt; green, 22 dpt) and late (right; blue, 34 dpt; red, 41 dpt; green, 48 dpt) time points. **d**, **e**, Density of spines (16 cells,  $n = 5$  mice) (**d**) and boutons (33 processes,  $n = 6$  mice) (**e**). Dark blue, average; light blue, individual cells/processes. **f**, **g**, Turnover (blue), fractional gain (green) and fractional loss (red) of spines (**f**) and boutons (**g**); increased turnover until 8 wpt (4–10 wpt,  $P < 0.0001$ , one-way repeated measure ANOVA, Tukey post-tests). **h**, **i**, Survival fraction of newly formed spines (**h**) and boutons (**i**) 4 dpt to 11 mpt. Spines and boutons formed early (red spectrum) are more prone to be eliminated. Structures formed at 9 mpt (blue) have a higher chance of survival (2,493 spines,  $n = 5$  mice; 1,600 boutons,  $n = 6$  mice; spines  $P < 0.00064$ , boutons  $P < 0.00018$ , Gehan–Breslow–Wilcoxon comparison, Bonferroni correction for multiple comparison). Grey denotes not significantly different. Scale bars, 10  $\mu$ m (**a**, **c**). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ .

### Efferent projections to correct target regions

To test whether transplanted neurons project to appropriate targets, red fluorescent protein (RFP)-expressing axons were examined throughout the brain at 5 wpt ( $n = 3$  mice; fibres in 38 anatomical regions; Fig. 3a–c). In V1, many axons were present in L5 (Fig. 3b, Vis), the prime laminar target of L2/3 neurons<sup>17</sup>. Ipsilateral visual areas contained the highest number of fibres, followed by the retrosplenial and ectorhinal cortices, and the corpus callosum (Fig. 3b, c), all of which are appropriate targets for V1 neurons. Axons also reached distant cortical regions via the white matter, for example, the motor and the contralateral visual cortex at 5 mm distance (Fig. 3b). We also observed projections to a few sub-cortical regions (cortical subplate regions and cerebral nuclei; Fig. 3b, c). A comparison with L2/3 V1 anterograde tracing data from the Allen Mouse Brain Connectivity Atlas (see Methods) shows that projections of transplanted neurons closely match the normal projectome (Fig. 3c, d). We did not detect axons in the thalamus and superior colliculus (Fig. 3c, d), both normally also not innervated by L2/3 V1 neurons. The few axons found in regions that are not regular targets of L2/3 V1 neurons (for example, amygdala) may be projections from transplanted neurons originating from other cortical areas (for example, orbital and cingulate regions), as we collected cells from the entire embryonic neocortex. We also detected axonal branching (Fig. 3b, insets; Fig. 3e) and, occasionally, RFP puncta near axons, assuming glial shape and aligned with interrupted fibres (Fig. 3e), suggesting that axonal pruning takes place by resident glia to refine projections.





**Figure 3 | Transplanted neurons extend long-range, largely cortical axonal projections.** **a**, Sagittal section, transplanted hemisphere shows high density of neurites in V1. **b**, **c**, Graft-derived projections at 5 wpt. See Extended Data Table 1 for abbreviations. **b**, Confocal images of the most densely innervated brain regions in both hemispheres. Schematics of the corresponding sagittal sections (Allen Mouse Brain Atlas, see Methods) indicate the medio-lateral position and depicted region (red rectangles). Red brackets denote insets magnified within respective or adjacent images (red arrowheads). DAPI denotes nuclear staining for anatomical reference; dashed and solid lines delimit anatomical regions and edge of sections, respectively. **c**, Colour-coded fibre density in various brain structures (dark to light blue indicates high to low density). Coloured bars, far right, represent main projections in naive mice, as shown in **d**. Empty boxes denote absent projections. **d**, Projections of Cux2<sup>+</sup> neurons in V1, categorized by main functional groups. Data from the Allen Mouse Brain Connectivity Atlas (see Methods). Note overrepresentation of cortical projections (isocortex, hippocampal formation (HPF)). Projections to thalamus (TH) and superior colliculus (SC) probably originate from a scarce population of Cux2<sup>+</sup> deep layer neurons. **e**, Putative phagocytosing glial cells (RFP puncta, empty arrowheads) following an interrupted fibre (solid-to-empty arrowheads). Scale bars, 100  $\mu$ m (**b**) and 50  $\mu$ m (**e**).

## Formation of appropriate afferent connectivity

To examine brain-wide synaptic inputs to transplanted neurons, we used modified rabies virus (RABV)-based monosynaptic tracing (see Methods)<sup>18</sup>. RFP-G-TVA-expressing neurons were grafted into lesioned V1, RABV (GFP) was injected into the transplantation site at 4 or 12 wpt, and labelled neurons were examined after 1 week (Fig. 4a, b). Starter neurons (RFP<sup>+</sup>/GFP<sup>+</sup>) were located exclusively within the transplant (Fig. 4c, d) and were surrounded by many GFP-only, monosynaptic input neurons (Fig. 4c). We found 20 anatomical regions containing cells innervating transplanted neurons at 4 wpt (Fig. 4e), with visual cortex exhibiting the highest connectivity ratio ( $21.50 \pm 5.07$ ,  $n = 6$  mice; Fig. 4f, Extended Data Fig. 5a–c), comparable to endogenous L2/3 neurons. Interareal connectivity included afferents from ipsilateral sensory, motor and association areas, as well as long-range connections from subcortical nuclei and contralateral cortex (Fig. 4e). Importantly, all of these regions are also known to project to V1 normally<sup>19,20</sup>, and we detected no aberrant input. We observed strong innervation from the dorsal lateral geniculate nucleus (dLGN), and from higher-order cortical areas, such as retrosplenial cortex (RS) and the posterior parietal association area (PPa). Connectivity ratios for these areas were, however, far lower than for local connections.

Remarkably, connectivity ratios were similar between transplanted and endogenous L2/3 neurons, which were transduced during development with the vector required for later monosynaptic tracing (Fig. 4f, Extended Data Fig. 6a–c). This was also true for areas with weak input, many of which could only be observed in mice with a sufficient number of starter neurons ( $>20$ , Extended Data Fig. 7a). Connectivity was similar for transplanted neurons at 4 and 12 wpt, with only a few additional input regions at 12 wpt, such as claustrum, pons and dorsal striatum, regions known to innervate the visual cortex<sup>20–22</sup> ( $n = 6$  mice per group; Fig. 4g, Extended Data Fig. 7b, c).

For proper circuit function, transplanted neurons in V1 must receive topographically organized inputs from the dLGN. We therefore determined the location of transplanted starter neurons in V1 and their respective inputs from the dLGN and found a clear spatial relationship, matching the normal mapping (Extended Data Fig. 8a–c, Supplementary Video 4). Notably, connectivity increased from 4 to 12 wpt, suggesting that further geniculocortical (graft) connections are added (Extended Data Fig. 8d). To quantify the topography further, we plotted the centroid of each cluster relative to the centroid of the respective anatomical structure (Extended Data Fig. 8e). Medio-lateral, antero-posterior and dorso-ventral/antero-posterior locations in the dLGN and V1, respectively, inversely correlate, matching the topographical organization of our control experiments (Extended Data Fig. 8e) and previous reports<sup>23,24</sup>. Thus, input to transplanted neurons resembles input to endogenous neurons at a quantitative level and includes topographically organized dLGN connections.

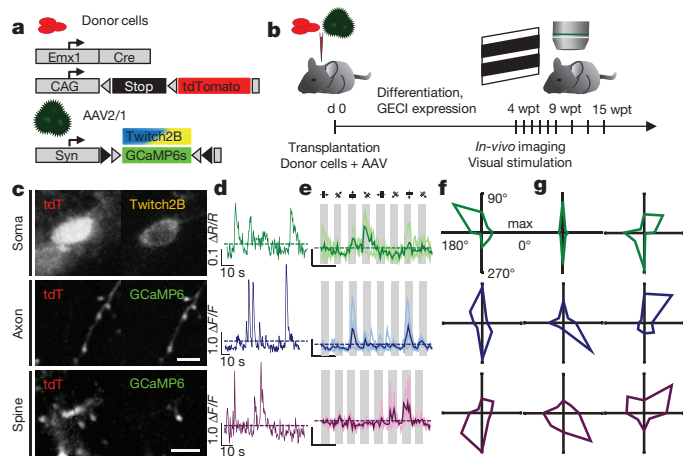
## Transplanted neurons process visual inputs

In V1, specific stimulus features, such as orientation and movement direction, are encoded at the level of single cells<sup>8</sup>. Thus, selective responses of transplanted neurons to such stimulus attributes would be highly indicative of their correct functional integration.

We transplanted embryonic neurons (tdTomato<sup>+</sup>) labelled with a genetically encoded calcium indicator (GECI) (such as GCaMP6s<sup>25</sup> or Twitch2B<sup>26</sup>; Fig. 5a). At 4–15 wpt, we presented moving gratings to the contralateral eye and recorded responses with two-photon imaging (Fig. 5b) by measuring changes in GECI fluorescence relative to baseline in somata, axons and dendritic spines (Fig. 5c–g, Extended Data Figs 9a and 10a). Almost all neurons with significant changes in fluorescence ( $P < 0.05$ , analysis of variance (ANOVA); Fig. 5d) exhibited stimulus-evoked responses ( $\Delta F/F > 3\sigma$  or  $\Delta R/R > 0.05$ , 27 out of 28 cells,  $n = 5$  mice; Fig. 5e), and approximately half of the visually responsive neurons showed strong orientation and/or direction preference (13 out of 27 cells,  $n = 5$  mice; Fig. 5f, g, Extended Data Fig. 10a), with sharp tuning properties typical for excitatory L2/3 neurons in adult mouse V1







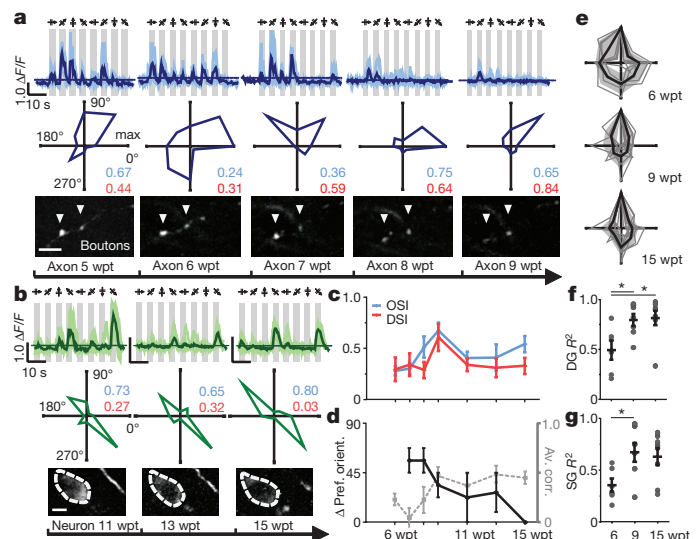
**Figure 5 | Transplanted neurons show tuned responses to visual stimuli.** **a**, Donor cells, co-labelled with Emx1-Cre-driven tdTomato and GECIs (GCaMP6s or Twitch2B; see Methods). **b**, Experimental timeline. **c–g**, Changes in GECI fluorescence in somata (green), axons (blue) and spines (magenta) of transplanted neurons. **c**, Examples of transplanted neurons (tdT<sup>+</sup>) expressing GECIs. Single optical section, maximum projection of all frames of one stimulation sequence. Scale bars, 5  $\mu$ m. **d**, Example traces of stimulus-evoked changes in fluorescence relative to baseline ( $\Delta R/R$ , Twitch2B;  $\Delta F/F$ , GCaMP6). **e**, Single (light traces) and average (dark traces) responses to stimulus presentations, sorted by grating orientation and direction. Grey, stimulus on; scale as in **d**. **f**, Normalized average maximum peak response at each direction plotted in polar coordinates for cells in **e**. **g**, Further examples of polar plots.

at 4 wpt were also found to provide input 2 months later. We conclude that transplanted neurons do not require prolonged periods for structural maturation or excessive re-wiring for correct incorporation into cortical circuits.

Inputs to transplanted neurons derive from regions that are all known to project to V1 (refs 19, 20), and they achieve quantitatively normal levels of monosynaptic input from most regions. This contrasts with a more limited connectivity observed previously<sup>7,34</sup>, probably owing to differences in the lesion model or donor cells. The few quantitative differences we found in connectivity compared to endogenous neurons may be due to the presence of transplanted neurons also in other layers (for example, L4, which receives denser input from the dLGN than L2/3; ref. 35). Also, since the contralateral cortex received multiple injections of beads, innervation from this partially damaged region is expected to be lower. In addition to precise brain-wide inputs, we found that geniculocortical (cortical) connections are topographically organized, an important prerequisite for proper visual processing. Thus, neuronal loss and subsequent transplantation do not lead to notable rewiring of dLGN axons within V1, which would cause aberrant connections and altered topography.

Functional imaging revealed initial broad tuning for orientation and direction at 4 wpt, despite high spine density and adequate inputs at this time. Tuning then sharpens substantially, probably reflecting activity-dependent functional refinement, driven by visual input<sup>36</sup>. Indeed, circuit remodelling is correlated with increased spine dynamics<sup>37</sup>, which we observe during this period. In line with this, monosynaptic tracing revealed that geniculate input increases from 1 to 3 mpt, supporting the notion of ongoing synapse refinement.

The slow sharpening of tuning observed in transplanted neurons differs from normal development, in which selective responses are present early, immediately after eye opening<sup>38</sup>, with adult-like values reached within only 1 week<sup>39</sup>. This initial, sharp tuning is thought to depend largely on feed-forward inputs conveying shared feature selectivity to clonally related neurons<sup>40,41</sup>. These are initially coupled by gap junctions that are subsequently replaced by conventional chemical synapses<sup>40</sup> through activity-dependent processes. Transplanted



**Figure 6 | Tuning of transplanted neurons sharpens over time.**

**a**, Changes in tuning at 5–9 wpt, measured from the axon of a transplanted cell (arrowheads: boutons). **b**, Stable tuning, transplanted cell at 11–15 wpt (dashed line, soma). **a, b**, Top row, individual and average stimulus evoked responses. Middle row, polar plots and orientation and direction selectivity indices (OSI/DSI). Bottom row, single plane maximum projections during visual stimulation. **c**, OSI/DSI at 6–15 wpt (27 cells,  $n = 5$  mice). **d**, Dark line, left y axis: difference in preferred orientation of neurons recorded in at least two subsequent time points from 6 to 15 wpt (14 out of 27 cells,  $n = 4$  out of 5 mice;  $P = 0.0491$ , Kruskal–Wallis test). Light dashed line, right y axis: correlation coefficients of trial-to-trial responses at the preferred direction of neurons recorded at least twice from 6 to 15 wpt (15 out of 27 cells,  $n = 5$  mice;  $P = 0.0482$ , Kruskal–Wallis test). **e**, Overlay of polar plots at 6, 9 and 15 wpt (maximum aligned to 90°). Overall sharpening of tuning (17 out of 27 cells,  $n = 4$  out of 5 mice, two-way ANOVA,  $P = 0.0029$ ). Grey, individual neurons; black, average; grey area,  $\pm$  s.d. **f, g**, Orientation and/or direction selectivity, Gaussian fits (Extended Data Fig. 11; Methods). Goodness-of-fit of double Gaussian (DG) (f) and single Gaussian (SG) (g) functions (17 out of 27 cells,  $n = 4$  out of 5 mice;  $P = 0.0100$  (DG) and  $P = 0.0382$  (SG), Kruskal–Wallis test, Dunn's post-test). Scale bars, 5  $\mu$ m (a, b). \* $P < 0.05$ .

neurons might have to skip this developmental step, and the formation of connections onto transplanted cells might be dominated by activity-dependent remodelling, starting from an initial state of low selectivity.

Taken together, we demonstrate here that transplanted embryonic neurons integrate into pre-existing cortical circuits challenged by an injury. They grow axons through the adult brain, reaching proper target areas, and receive V1-specific inputs from host neurons. Notably, after 2–3 months, the transplanted neurons are fully integrated, showing functional properties that are indistinguishable from their original counterparts. Thus, neuronal replacement therapies may be realistic, at least at times when a sufficient part of the pre-existing neuronal network is still available.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 15 December 2015; accepted 23 September 2016.**

**Published online 26 October 2016.**

- Lindvall, O., Kokaia, Z. & Martinez-Serrano, A. Stem cell therapy for human neurodegenerative disorders-how to make it work. *Nat. Med.* **10** (suppl.), S42–S50 (2004).
- Thompson, L. H. & Björklund, A. Reconstruction of brain circuitry by neural transplants generated from pluripotent stem cells. *Neurobiol. Dis.* **79**, 28–40 (2015).
- Barker, R. A., Barrett, J., Mason, S. L. & Björklund, A. Fetal dopaminergic transplantation trials and the future of neural grafting in Parkinson's disease. *Lancet Neurol.* **12**, 84–91 (2013).

4. Grealish, S. *et al.* Monosynaptic tracing using modified rabies virus reveals early and extensive circuit integration of human embryonic stem cell-derived neurons. *Stem Cell Reports* **4**, 975–983 (2015).
5. Shin, J. J. *et al.* Transplanted neuroblasts differentiate appropriately into projection neurons with correct neurotransmitter and receptor phenotype in neocortex undergoing targeted projection neuron degeneration. *J. Neurosci.* **20**, 7404–7416 (2000).
6. Cunningham, M. *et al.* hPSC-derived maturing GABAergic interneurons ameliorate seizures and abnormal behavior in epileptic mice. *Cell Stem Cell* **15**, 559–573 (2014).
7. Michelsen, K. A. *et al.* Area-specific reestablishment of damaged circuits in the adult cerebral cortex by cortical neurons derived from mouse embryonic stem cells. *Neuron* **85**, 982–997 (2015).
8. Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* **28**, 7520–7536 (2008).
9. Madison, R. D. & Macklis, J. D. Noninvasively induced degeneration of neocortical pyramidal neurons *in vivo*: selective targeting by laser activation of retrogradely transported photolytic chromophore. *Exp. Neurol.* **121**, 153–159 (1993).
10. Miller, M. Maturation of rat visual cortex. I. A quantitative study of Golgi-impregnated pyramidal neurons. *J. Neurocytol.* **10**, 859–878 (1981).
11. Alvarez-Dolado, M. *et al.* Fusion of bone-marrow-derived cells with Purkinje neurons, cardiomyocytes and hepatocytes. *Nature* **425**, 968–973 (2003).
12. Brilli, E. *et al.* Neural stem cells engrafted in the adult brain fuse with endogenous neurons. *Stem Cells Dev.* **22**, 538–547 (2013).
13. Gorski, J. A. *et al.* Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage. *J. Neurosci.* **22**, 6309–6314 (2002).
14. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
15. Holtmaat, A. *et al.* Long-term, high-resolution imaging in the mouse neocortex through a chronic cranial window. *Nat. Protocols* **4**, 1128–1144 (2009).
16. Hofer, S. B., Mrsic-Flogel, T. D., Bonhoeffer, T. & Hübener, M. Experience leaves a lasting structural trace in cortical circuits. *Nature* **457**, 313–317 (2009).
17. Xu, X. *et al.* Primary visual cortex shows laminar-specific and balanced circuit organization of excitatory and inhibitory synaptic connectivity. *J. Physiol. (Lond.)* **594**, 1891–1910 (2016).
18. Wickersham, I. R. *et al.* Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. *Neuron* **53**, 639–647 (2007).
19. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).
20. Zingg, B. *et al.* Neural networks of the mouse neocortex. *Cell* **156**, 1096–1111 (2014).
21. Waterhouse, B. D., Border, B., Wahl, L. & Mihailoff, G. A. Topographic organization of rat locus coeruleus and dorsal raphe nuclei: distribution of cells projecting to visual system structures. *J. Comp. Neurol.* **336**, 345–361 (1993).
22. Royce, G. J. & Laine, E. J. Efferent connections of the caudate nucleus, including cortical projections of the striatum and other basal ganglia: an autoradiographic and horseradish peroxidase investigation in the cat. *J. Comp. Neurol.* **226**, 28–49 (1984).
23. Pfeifferberger, C., Yamada, J. & Feldheim, D. A. Ephrin-As and patterned retinal activity act together in the development of topographic maps in the primary visual system. *J. Neurosci.* **26**, 12873–12884 (2006).
24. Schuett, S., Bonhoeffer, T. & Hübener, M. Mapping retinotopic structure in mouse visual cortex with optical imaging. *J. Neurosci.* **22**, 6549–6559 (2002).
25. Chen, T. W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
26. Thestrup, T. *et al.* Optimized ratiometric calcium sensors for functional *in vivo* imaging of neurons and T lymphocytes. *Nat. Methods* **11**, 175–182 (2014).
27. Ko, H., Mrsic-Flogel, T. D. & Hofer, S. B. Emergence of feature-specific connectivity in cortical microcircuits in the absence of visual experience. *J. Neurosci.* **34**, 9812–9816 (2014).
28. Marshel, J. H., Garrett, M. E., Nauhaus, I. & Callaway, E. M. Functional specialization of seven mouse visual cortical areas. *Neuron* **72**, 1040–1054 (2011).
29. Lewis, T. L., Jr, Courchet, J. & Polleux, F. Cell biology in neuroscience: Cellular and molecular mechanisms underlying axon formation, growth, and branching. *J. Cell Biol.* **202**, 837–848 (2013).
30. Adelson, J. D. *et al.* Developmental sculpting of intracortical circuits by MHC class I H2-Db and H2-Kb. *Cereb. Cortex* **26**, 1453–1463 (2014).
31. Cruz-Martín, A., Crespo, M. & Portera-Cailliau, C. Delayed stabilization of dendritic spines in fragile X mice. *J. Neurosci.* **30**, 7793–7803 (2010).
32. Fricker-Gates, R. A., Shin, J. J., Tai, C. C., Catapano, L. A. & Macklis, J. D. Late-stage immature neocortical neurons reconstruct interhemispheric connections and form synaptic contacts with increased efficiency in adult mouse cortex undergoing targeted neurodegeneration. *J. Neurosci.* **22**, 4045–4056 (2002).
33. Gaillard, A. *et al.* Anatomical and functional reconstruction of the nigrostriatal pathway by intranigral transplants. *Neurobiol. Dis.* **35**, 477–488 (2009).
34. Gaillard, F., Girman, S. V. & Gaillard, A. Afferents to visually responsive grafts of embryonic occipital neocortex tissue implanted into V1 (Oc1) cortical area of adult rats. *Restor. Neurol. Neurosci.* **12**, 13–25 (1998).
35. Alonso, J. M. & Swadlow, H. A. Thalamocortical specificity and the synthesis of sensory cortical receptive fields. *J. Neurophysiol.* **94**, 26–32 (2005).
36. Kreile, A. K., Bonhoeffer, T. & Hübener, M. Altered visual experience induces instructive changes of orientation preference in mouse visual cortex. *J. Neurosci.* **31**, 13911–13920 (2011).
37. Keck, T. *et al.* Massive restructuring of neuronal circuits during functional reorganization of adult visual cortex. *Nat. Neurosci.* **11**, 1162–1167 (2008).
38. Rochefort, N. L. *et al.* Development of direction selectivity in mouse cortical neurons. *Neuron* **71**, 425–432 (2011).
39. Hoy, J. L. & Niell, C. M. Layer-specific refinement of visual cortex function after eye opening in the awake mouse. *J. Neurosci.* **35**, 3370–3383 (2015).
40. Yu, Y. C. *et al.* Preferential electrical coupling regulates neocortical lineage-dependent microcircuit assembly. *Nature* **486**, 113–117 (2012).
41. Li, Y. *et al.* Clonally related visual cortical neurons show similar stimulus feature selectivity. *Nature* **486**, 118–121 (2012).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank D. Franzen, G. Jäger, T. Simon, V. Staiger, H. Tultschin and F. Voss for technical support, and A. Lepier for viral vector expertise. M. Sperling, P. Goltstein and A. Grade helped with hardware and software. This work was supported by the German Research Foundation (SFB 870 ‘Neuronal Circuits’: M.G., L.D., K.-K.C., T.B. and M.H.; SPP 1757: M.G. and L.D.), the Advanced ERC grant ChroNeuroRepair (M.G.), the Helmholtz Alliance Icmem (M.G.), the Boehringer Ingelheim Fonds (S.F.), and the Max Planck Society (S.F., T.B. and M.H.).

**Author Contributions** The original idea for the study came from M.G. M.G., M.H. and T.B. then initiated the study and planned the experimental approach. S.F., S.G., M.G., M.H. and L.D. designed the experiments; S.F. performed *in vivo* imaging experiments and analysis; S.G. performed the experiments in fixed tissue including connectivity experiments and analysis. M.G. and L.D. provided the lesion model. K.-K.C. provided the rabies virus and expertise for its use for monosynaptic tracing. Finally, M.G., M.H., S.G. and S.F. wrote the paper with input from T.B., L.D. and K.-K.C.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.G. (magdalena.goetz@helmholtz-muenchen.de) or M.H. (mark@neuro.mpg.de).



## METHODS

**Mice and anaesthetics.** All animal experiments were carried out in compliance with the institutional guidelines of the Max Planck Society, the Ludwig-Maximilians-University and the local government (Government of Upper Bavaria).

Data for this study are derived from a total of 38 adult mice of both sexes. For the characterization of the lesion model, the cells' identity and the afferent/efferent connectivity we used C57BL/6J mice. For cell fusion controls and structural *in vivo* two-photon imaging we used Emx1-Cre<sup>13</sup> crossed to CAG-CAT-GFP<sup>42</sup> donor cells and Ai9 (Rosa-CAG-LSL-tdTomato reporter mice)<sup>14</sup> host mice. For functional *in vivo* two-photon imaging we used Emx1-Cre<sup>13</sup> crossed to Ai9 donor cells and C57BL/6J host mice. No a priori determination of sample size was conducted, experiments were not randomized, and investigators were not blinded to experimental conditions, except where noted.

Mice were housed in a 12:12h light–dark cycle. All mice were 2–4 months old at the time of the first surgery. Surgeries were performed aseptically under anaesthesia with a mixture of fentanyl (0.05 mg kg<sup>-1</sup>, Hexal or Janssen), midazolam (5 mg kg<sup>-1</sup>, Ratiopharm or Roche) and medetomidine (0.5 mg kg<sup>-1</sup>, Orion Corp. or Fort Dodge). After surgery, anaesthesia was terminated with atipamezol (2.5 mg kg<sup>-1</sup>, Orion Corp.), flumazenil (0.5 mg kg<sup>-1</sup>, Hexal) and naloxone (1.2 mg kg<sup>-1</sup>, Ratiopharm) or buprenorphin (0.1 mg kg<sup>-1</sup>, Essex). Carprofen (5 mg kg<sup>-1</sup>, Pfizer) or meloxicam (1 mg kg<sup>-1</sup>, Boehringer Ingelheim) was administered as a postoperative analgesic.

**Neocortical layer-specific lesion.** Rhodamine-labelled latex beads (Lumafluor) were conjugated with chlorine e6 (Ce6, Frontier Scientific) as described previously<sup>43</sup>, with minor modifications. In brief, 1.5 ml of Ce6 solution (0.597 mg ml<sup>-1</sup> in 0.01 M phosphate buffer (PB), pH 7.4) were added to 5 mg of 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (MP Biomedicals) and kept at 4°C for 30 min. Then 12.5 µl of rhodamine latex beads were diluted in 100 µl PB, mixed with 750 µl of activated Ce6 solution, and incubated on an orbital shaker for 1 h at room temperature. The reaction was stopped by adding 335 µl of 0.1 M glycine buffer (pH 8.0). The conjugated latex beads were washed at least three times (30 min, 100,000g) with 0.01 M PB and finally resuspended in 50 µl PB. Beads were stored at 4°C and used within 1 month. For beads injection, mice were anaesthetized, and a unilateral craniotomy (2.5 mm diameter) was performed at the posterior end of the parietal bone, centred at 2.5 mm lateral from lambda, to expose the visual cortex. Ce6-conjugated beads (9 µg µl<sup>-1</sup>) were injected at 5–10 locations to reach a total volume of 0.5–1 µl, at a depth of 0.5 mm into the primary visual cortex (V1). Injections were restricted to the binocular zone of V1 with a total volume of 0.5 µl in structural and functional *in vivo* imaging experiments. The location of V1 and the binocular zone was identified using intrinsic optical imaging<sup>16</sup> and/or the characteristic blood vessel pattern. One to two weeks after the beads injection, a second craniotomy was performed in order to expose the contralateral V1 that was non-invasively subjected to laser-photoactivation of Ce6 for 4 min (670 nm, 27 mW; laser and beam shaping optics for a collimated parallel beam of 1 mm diameter, Schäfer+Kirchhoff) to induce apoptotic cell death of callosal projection neurons (CPNs). This lesion model thus targets only a fraction of neurons in an area and layer-specific manner, while the overall tissue structure is preserved.

**In utero electroporation.** For monosynaptic rabies virus tracing experiments, mouse cortical cells were labelled via *in utero* electroporation<sup>44</sup>. In brief, at embryonic day E14.5–E16.5, pregnant mice were anaesthetized and the uterine horns were exposed by laparotomy. The plasmid pDsRedExpress2-2A-Glyco-IRES2-TVA or pDsRedExpress2-IRES2-TVA (1 µg µl<sup>-1</sup>) was mixed with 2.5 µg µl<sup>-1</sup> Fast Green dye (Sigma-Aldrich) and 1 µl of the solution was injected into one of the lateral ventricles of each embryo using a glass capillary. Tweezer-type circular electrodes oriented with a ~45° angle for cortical targeting were used to hold the head of the embryo and deliver electrical pulses (5 pulses, 35 V, 100 ms) produced by a square-wave electroporation generator (ECM 830, BTX, Harvard Apparatus). Uterine horns were returned into the abdominal cavity and embryos were allowed to continue their normal development. For transplantation, E18.5 embryos previously electroporated at E14.5 were euthanized and cerebral cortices were collected for cell dissociation. Some cells were plated directly after dissociation at a density of 500,000 cells per well in 24-well plates in B27-containing DMEM high glucose (4.5 g l<sup>-1</sup>) with glutamax, plus penicillin–streptomycin, and fixed after 2 h for subsequent immunocytochemistry (see below).

For tracing of the afferents to endogenous upper layer neurons, electroporation was performed at E15.5–E16.5, and animals were allowed to develop until adulthood, when the RABV was then injected into V1 of the electroporated hemisphere.

**Embryonic cell culture and viral labelling.** In a subset of experiments, *in vitro* transduction with Moloney murine leukaemia virus (MMuLV)-derived retroviral vectors (*n* = 10 host mice) or adeno-associated virus (AAV; *n* = 2 host mice) was performed to label donor cells.

In brief, neocortical tissue from E14.5/E15.5 embryos was mechanically dissociated in HBSS (Invitrogen) buffered with 10 mM HEPES, or in EBSS with papain

(20 U ml<sup>-1</sup>, 0.005% DNase, 1 mM L-cysteine, 0.5 mM EDTA; Papain dissociation system, Worthington) after 45 min of enzymatic treatment at 37°C. Enzymatic activity was stopped by protease inhibitors (10 mg ml<sup>-1</sup> ovomucoid). Cells were plated in 20 µg ml<sup>-1</sup> poly-D-lysine (Sigma-Aldrich)-coated 24-well plates, at a density of 300,000 cells per well. Cells were initially kept in 10% FBS-containing DMEM high glucose (4.5 g l<sup>-1</sup>) with glutamax, plus penicillin–streptomycin (all from Invitrogen), and transduced after 2–4 h with viral vectors (1–2 µl per well; titres typically ranged from 10<sup>7</sup> to 10<sup>11</sup> transducing units ml<sup>-1</sup>). For monosynaptic rabies tracing experiments MMuLV-derived retroviral vectors (CAG-DsRedExpress2-2A-Glyco-IRES2-TVA or CAG-DsRedExpress2-IRES2-TVA) and for functional *in vivo* imaging experiments an AAV vector (AAV2/1-hSyn1-flex-mRuby2-P2A-CGAMP6 s-WPRE-SV40) was used. Serum was gradually removed by replacing half of the medium with B27-containing DMEM high glucose (4.5 g l<sup>-1</sup>) with glutamax plus penicillin–streptomycin on each of the following 2 days. Cells were collected for transplantation after 2–5 days *in vitro*.

**Transplantation of embryonic cells into V1.** Embryonic cells for transplantation were fluorescently labelled either in the above mentioned mouse lines, by *in utero* electroporation of DNA plasmids, or via *in vitro* viral transduction.

Seven to ten days after laser-photoactivation, embryonic cells were transplanted into the previously illuminated area. In short, E18.5 embryos were euthanized, and fluorescently labelled cortical hemispheres were collected for dissociation. Cortical tissue was mechanically dissociated either in HBSS buffered with 10 mM HEPES, or in EBSS with papain (20 U ml<sup>-1</sup> papain, 0.005% DNase, 1 mM L-cysteine, 0.5 mM EDTA; Papain dissociation system, Worthington) after 45 min of enzymatic treatment at 37°C. In case of papain treatment, enzyme activity was stopped by protease inhibitors (10 mg ml<sup>-1</sup> ovomucoid). A cell suspension (50 million cells ml<sup>-1</sup>) was prepared in neurobasal medium (NB) or DMEM high glucose (4.5 g l<sup>-1</sup>), both supplemented with B27, glutamax and penicillin–streptomycin.

Donor cells labelled via *in vitro* viral transduction were washed at least 5 times with pre-warmed PBS to remove any remaining viral particles. Gentle trypsinization (0.025%, 10 min at 37°C) was performed and a cell suspension (25 million–50 million cells ml<sup>-1</sup>) was prepared in B27-containing DMEM high glucose (4.5 g l<sup>-1</sup>) with glutamax, plus penicillin–streptomycin.

Between 25,000 and 80,000 donor cells were transplanted into V1 of adult mice. A total volume of 1 µl of cell suspension was injected with a syringe (gauge 31–33, Hamilton) within the area previously illuminated to induce neuronal death, at 2.5 ± 0.2 mm medio-lateral, 0.0 ± 0.2 mm antero-posterior relative to lambda, and distributed dorso-ventrally over 200–300 µm at a depth between 0.5 and 0.2 mm. The exact injection coordinates and pattern of pial vasculature were noted for later injection of the RABV. For two-photon imaging a cranial glass window was implanted, otherwise the bone lid was repositioned and the skin was sutured.

**Cranial window implant.** For chronic structural and functional *in vivo* imaging experiments a cranial glass window was implanted on top of V1 (ref. 15) after cell transplantation. In brief, a coverslip (5 mm diameter, #1 thickness, EMS) was loosely placed on the dura, resting on the edge of the craniotomy, and sealed to the bone with cyanoacrylate. A small metal bar with screw holes (5 × 8 mm) for head fixation during image acquisition was attached to the skull medial to the window implant. Skin margins, cover-glass and metal bar were embedded in dental acrylic (Heraeus Kulzer) mixed with black pigment (Kremer).

**Intrinsic optical imaging.** In a subset of mice (*n* = 16), the location of cell transplantation was verified with optical imaging of intrinsic signals. In short, anaesthetized mice were presented with square wave drifting gratings (4 orientations, 600 ms stimulus duration; 0.03 cycles degree<sup>-1</sup>, 2 cycles s<sup>-1</sup>) in a 2 × 2 array covering approximately –10° to 70° azimuth, –20° to 40° elevation of the ipsi- or contralateral visual field, respectively. For identification of the binocular zone, visual stimuli were presented to the ipsilateral eye while the contralateral eye was covered. The cortical surface was evenly illuminated through the cranial window with monochromatic light at 707 nm. A cooled, slow scan CCD camera (12 bit, Optical Imaging Inc.) was focused 200–300 µm below the cortical surface and frames were recorded with 600 ms exposure time. Images of average responses (3 repetitions of 12 stimulus frames per location) were blank-corrected, range-fitted and low-pass-filtered<sup>24</sup>. The false colour-coded maximum projection of visual responses is mapped on top of the blood vessel image acquired before visual stimulation.

**Structural *in vivo* two-photon imaging.** *In vivo* two-photon imaging was carried out on an Olympus FV1000BX61 system equipped with a mode-locked Ti:sapphire laser (Mai Tai DeepSee, Spectra-Physics) through a 25× water immersion objective (1.05NA, Olympus). Laser settings and image acquisition were controlled by Fluoview software (Olympus).

For structural *in vivo* imaging mice were anaesthetized with fentanyl based anaesthesia (see above) and placed on a feedback controlled heating pad. Data were acquired at 910 nm with an average laser power of <30 mW below the objective, and the emission signal was directed through a dichroic mirror (DM570) and red/green bandpass emission filters (BA495–540HQ and BA570–625HQ, all Olympus);

a typical imaging session lasted 2 h. Host mice ( $n = 11$ ) were imaged as early as 3 dpt. Individual transplanted cells were identified and followed in short, increasing intervals (2–5 days) within the first 4 weeks and weekly thereafter (up to 12 wpt;  $n = 6$  mice). In two mice we acquired late time points at 9–11 mpt.

In each imaging session, high-resolution tiled volume stacks ( $510 \times 510 \mu\text{m}$  field of view;  $0.33 \mu\text{m}$  per pixel;  $1\text{--}3 \mu\text{m}$  z-steps) were acquired up to a depth of  $350 \mu\text{m}$  from the pial surface for overview and reconstruction of whole cell morphology of transplanted cells. In addition, high-resolution close-up stacks ( $73 \times 73 \mu\text{m}$ ;  $0.14 \mu\text{m}$  per pixel;  $0.5\text{--}1.0 \mu\text{m}$  z-steps) of typically three individual dendritic and axonal processes were acquired at various depths between 50 and  $300 \mu\text{m}$ . For the analysis presented in Fig. 2, we included both basal and apical dendritic processes.

**In vivo two-photon calcium imaging.** For functional *in vivo* imaging experiments, donor cells were labelled with the GECIs CGaMP6s<sup>25</sup> or Twitch2B<sup>26</sup> (a FRET-based sensor). Emx1-Cre  $\times$  Ai9 donor cells were mixed with AAV encoding a double-floxed inverted open reading frame version of either GECI before transplantation (AAV2/1-hSyn1-flex-CGaMP6s-WPRE-SV40; AAV2/1-CAG-flex-Twitch2B-WPRE-SV40). In a subset of experiments ( $n = 2$  host mice) donor cells were labelled *in vitro* (AAV2/1-hSyn1-flex-mRuby2-P2A-CGaMP6s-WPRE-SV40, see above).

*In vivo* imaging was performed under light anaesthesia; mice received  $0.4 \times$  dose for initial anaesthesia and a subsequent  $0.2 \times$  dose every 90 min. Mice were kept on a feedback-controlled heating pad. For ipsi- and contralateral visual stimulation, either the left or the right eye was occluded, respectively, and full field moving gratings (square wave, high contrast;  $0.04$  cycles degree<sup>-1</sup>,  $1.5$  cycles s<sup>-1</sup>; 4 orientations, 8 directions) were presented to the open eye (30 cm distance monitor to eye). The 8 directions were presented in random order, each displayed for 3 s, followed by 3 s of an isoluminant grey screen. Presentation of  $3 \times 8$  directions was flanked by 10 s of grey screen (constituting one stimulus sequence of 3 repeats). Typically, 2–3 stimulus sequences were presented per imaging plane (altogether 6–9 repeats).

Data were acquired either at 940 nm (GCaMP6) or at 860 nm (Twitch2B) with an average laser power of  $<30$  mW; a typical imaging session lasted 2–3 h. Emitted light was directed through a longpass dichroic mirror (DM570, GCaMP6; 505DCXR, Twitch2B) and recorded through emission filters (BA495–540HQ and BA570–625HQ, GCaMP6; ET480/40M and ET535/30M, Twitch2B).

In each imaging session a tiled volume stack ( $760 \times 760 \mu\text{m}$  field of view;  $0.49 \mu\text{m}$  per pixel;  $3\text{--}5 \mu\text{m}$  z-steps) was acquired up to a depth of  $350 \mu\text{m}$  from the pial surface for an overview of transplanted (RFP<sup>+</sup>) neurons. Five to ten candidate areas with RFP<sup>+</sup>/GECI<sup>+</sup> neurons were recorded at a frame rate of 2.4 Hz during visual stimulation ( $73 \times 73 \mu\text{m}$  field of view;  $0.28 \mu\text{m}$  per pixel).

Host mice were imaged starting at 4 wpt, and individual responsive neurons were repeatedly recorded up to 15 wpt. In a subset of experiments ( $n = 2$  mice, see above), we recorded late time points at 11–15 wpt.

**Analysis of *in vivo* structural data.** Image stacks were processed using the Fiji package of ImageJ (US National Institutes of Health) as follows. Fluorescence signals of rhodamine-coupled Ce6 beads were removed by channel subtraction. Images were converted to 8-bit and subjected to a small 2D Gaussian filter ( $\sigma < 0.6$  pixels). For display purposes only, maximum intensity z-projections are shown with adjusted brightness/contrast.

Whole-cell morphology was reconstructed using Simple Neurite Tracer. In brief, apical and basal dendrites were semi-automatically traced through the high-resolution tiled volume stack. On the basis of the traced skeleton, a single-cell 3D volume model was rendered.

To determine spine and bouton densities, dynamics and survival, putative synaptic structures were identified<sup>15</sup> in image stacks at each recorded time point. We included all clearly visible structures in  $x$ ,  $y$  and  $z$ . An ID was assigned to each individual identified structure at the time point of its first appearance and registered across time points. Initial identification of spines and boutons was performed blinded to the time point of recording. Registration across time points was performed sequentially. Density is reported as structures per  $\mu\text{m}$ , and turnover is calculated as fraction of structures (gained + lost)/(total  $t1 + t2$ ). We calculated the average survival fraction of gained structures dependent on the time point of their first appearance (newly formed structures at binned time points: 3–4, 5–6, 7–9, 12–13, 17–19 and 22–24 dpt and weekly bins between 4 and 9 wpt) according to the non-parametric Kaplan–Meier estimator (using GraphPad PRISM). This method takes into account some uncertainty of the actual survival of spines present at (and presumably longer than) the last experimental time point. Hazard ratios compare the rate of structure loss between structures that were gained at different time points; median survival ratios compare the relative median survival of gained structures at different time points.

In total, 16 dendritic stretches from 5 mice, and 33 axonal stretches from 6 mice were analysed (average dendritic length:  $50.5 \pm 12.4 \mu\text{m}$ ; average axonal length:  $77.7 \pm 25.3 \mu\text{m}$ ). A total of 13,251 dendritic spines and 6,266 axonal boutons across all time points were identified and registered to 2,493 individual dendritic spines and 1,600 individual axonal boutons on  $0.8$  mm total dendritic and  $2.6$  mm total axonal length.

**Analysis of *in vivo* functional data.** Functional imaging data were analysed using Fiji and Matlab (Mathworks), and the investigator was blinded to the recording time points analysed. Individual frames were background subtracted using a rolling ball algorithm ( $>100$  pixels radius), converted to 8 bit and subjected to a small 2D Gaussian filter ( $\sigma < 0.8$  pixels). Stacks were full-frame aligned using linear transformations (StackReg, P. Thévenaz, EPFL; <http://bigwww.epfl.ch/thevenaz/stackreg>) and regions of interest were selected manually based on the aligned maximum intensity projection across a stimulation sequence. The fluorescence signal ( $F$ ) was calculated as the average fluorescence of all pixels within a given region of interest (Twitch2B:  $R = \text{av.}F_{\text{YFP}}/\text{av.}F_{\text{CFP}}$ ). Neuronal activity was measured as the normalized change in fluorescence signal over time:  $(F_t - F_0)/F_0$  (Twitch2B:  $R_t - R_0/R_0$ ). The baseline ( $F_0$ ,  $R_0$ ) was calculated as the average signal over typically 10 s before and after each stimulation sequence (see above).

We classified neurons as visually responsive if the average  $\Delta F/F_0 > 3\sigma$  (Twitch2B:  $\Delta R/R_0 > 0.05$ ; ref. 33) for at least one stimulus direction. Tuning properties of each neuron were depicted in complex space using the normalized average peak response for each direction expressed in polar coordinates.

Orientation and direction tuning preference was expressed as orientation and direction selectivity index (OSI and DSI), respectively, and calculated as follows:  $\text{OSI} = (R_{\text{pref}} - R_{\text{ortho}})/(R_{\text{pref}} + R_{\text{ortho}})$ ; and  $\text{DSI} = (R_{\text{pref}} - R_{\text{opposite}})/(R_{\text{pref}} + R_{\text{opposite}})$ .  $R$  is the average peak response to the preferred direction ( $R_{\text{pref}}$ ), to the mean of the orthogonal directions ( $R_{\text{ortho}}$ ) and the opposing direction ( $R_{\text{opposite}}$ ). As ratio based tuning properties do not take into account the distribution of responses across all tested directions, we also calculated single and double Gaussian fits<sup>45</sup>. Following the assumption that an ideal orientation (or direction) tuned neuron would be perfectly described by a double (or single) Gaussian fit, the goodness of fit ( $R^2$ ) serves as a measure of tuning quality. In short, curves were fit with non-linear regression using PRISM (GraphPad). Single Gaussian fits were calculated according to  $y = a + \text{amp} \times \exp(-0.5 \times ((x - x_{\text{mean}})/\sigma)^2)$ , with  $a$  = offset from  $x$ -axis,  $\text{amp}$  = peak amplitude,  $x$  = stimulus directions in degrees,  $x_{\text{mean}} = x$  value at peak amplitude, and the following constraints:  $a > 0$ ,  $\text{amp} = 0\text{--}1$ ,  $x_{\text{mean}} = 180^\circ$ ,  $\sigma \geq 22.5^\circ$ . Double Gaussian fits were calculated according to  $y = a + \text{amp1} \times \exp(-0.5 \times ((x - x_{\text{mean1}})/\sigma_1)^2) + \text{amp2} \times \exp(-0.5 \times ((x - x_{\text{mean2}})/\sigma_2)^2)$ , with  $a$  = offset from  $x$ -axis,  $\text{amp1}$  = peak amplitude,  $\text{amp2}$  = amplitude at opposing direction,  $x$  = stimulus directions in degree,  $x_{\text{mean1}} = x$  value at peak amplitude,  $x_{\text{mean2}} = x$  value at amplitude of opposing direction, and the following constraints:  $a > 0$ ,  $\text{amp1} = 0\text{--}1$ ,  $\text{amp2} < \text{amp1}$ ,  $\Delta x_{\text{mean1,2}} = 180^\circ$ ,  $\sigma_{1,2} \geq 22.5^\circ$ .  $R^2$  is computed as the normalized sum of least squares.

To describe further the changes in tuning of individual neurons over time, we calculated the difference in preferred orientation between successive imaging time points for all neurons recorded at least twice without interruption (14 out of 27 cells,  $n = 4$  mice). With 4 orientations, the individual difference was limited to discrete values of  $\Delta 45^\circ$  between  $0^\circ$  and  $90^\circ$ . Note, that changes in preference and specificity reported here are unlikely to arise from nonlinearities of the calcium indicators, as we see both, marked changes in tuning preference at consistent average peak amplitudes across time points (Fig. 6d, Extended Data Fig. 11e, f), as well as stable tuning at late time points (Fig. 6b).

Changes in the reliability of responses over time were assessed for all neurons recorded at least twice (15 out of 27 cells,  $n = 5$  mice) by calculating the average correlation coefficients (Pearson correlation) of trial-to-trial responses during visual stimulation at the preferred direction across time points. As an additional measure of reliability, we calculated the average coefficient of variation ( $\text{CV} = \sigma/\mu$ ) for GCaMP6s<sup>+</sup> cells recorded at 6, 9 and 15 wpt (15 out of 27 cells,  $n = 3$  mice; 5–7 cells per time point). Note that owing to the distinct ranges of peak amplitudes of GCaMP6s and Twitch2B it is not possible to calculate a meaningful CV value combining cells labelled with either GECI, and thus only GCaMP6s<sup>+</sup> cells were included. **Monosynaptic rabies virus tracing.** We examined the brain-wide synaptic input to transplanted cells by retrograde monosynaptic tracing with a modified rabies virus (EnvA- $\Delta$ G-RABV)<sup>18</sup>. RABV particles are pseudotyped to specifically infect cells expressing TVA-class receptors and express eGFP instead of their own glycoprotein (G). Transsynaptic spread to pre-synaptic partners is only possible after complementing the RABV with its G-protein coat by the G-TVA expressing cells and is therefore limited to one monosynaptic jump. G-TVA expressing cells also encode for DsRed, while their pre-synaptic partners will be eGFP<sup>+</sup> only. DsRed<sup>+</sup>/GFP<sup>+</sup> neurons are termed starter cells.

In brief, EnvA- $\Delta$ G-RABV was injected 4 or 12 weeks after transplantation of embryonic cells expressing DsRedExpress2-2A-Glyco-IRES2-TVA or DsRedExpress2-IRES2-TVA, in three locations surrounding the transplantation site (200 nl per location). To map the pre-synaptic connectivity of upper layer neurons in adult V1, E14.5–E16.5 embryos were *in utero* electroporated with pDsRedExpress2-2A-Glyco-IRES2-TVA or pDsRedExpress2-IRES2-TVA (see above). Embryos were allowed to develop and as adult mice received 3 injections of RABV within V1 of the electroporated hemisphere (200 nl per injection). In both



experimental groups, animals were euthanized 7–9 days later for immunostainings and circuit analysis.

**Immunocyto/histochemistry.** Plated cells were fixed in 4% paraformaldehyde (PFA) for 30 min, washed and incubated in blocking and permeabilizing solution for 30 min (3% bovine serum albumin; 0.5% Triton X-100) before applying the primary antibodies goat mCherry (1:500; Sicgen), rabbit anti-Cux1 (1:500; Santa Cruz), for overnight incubation at 4 °C. After washing, cells were incubated with species- and subclass-specific secondary antibodies conjugated to Cy3 and Cy5 (Dianova) used at 1:500 for 2 h at room temperature. Nuclei were stained with  $1\ \mu\text{g ml}^{-1}$  4,6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich; 5 min at room temperature) and coverslips were mounted on glass slides with Aqua-Poly/Mount (Polysciences).

For immunohistochemistry, brains were collected after transcardial perfusion of deeply anaesthetized mice with PBS (5 min) followed by 4% PFA, 30–40 min. Brains were then post-fixed in 4% PFA overnight, at 4 °C, serially cut on a vibratome into 60–70  $\mu\text{m}$  sagittal or coronal sections, which were kept free-floating for further processing. TUNEL staining was performed according to the manufacturer's instructions (Roche). Immunohistochemistry was carried out using the following primary antibodies: chicken anti-GFP (1:1,000; Aves Labs), rabbit anti-RFP (1:1,000; Rockland), goat anti-mCherry (1:200; Sicgen), rabbit anti-Cux1 (1:200; Santa Cruz). After washing, sections were incubated with species- and subclass-specific secondary antibodies conjugated to Cy3 or Cy5 (Dianova) or Alexa Fluor 488 or 647 (Invitrogen), used at 1:500 or 1:1,000 depending on high ( $\geq 1:500$ ) or low ( $< 1:500$ ) concentration of the primary antibody. Sections were incubated for 10 min with  $1\ \mu\text{g ml}^{-1}$  DAPI for nuclear labelling and mounted on glass slides with Aqua-Poly/Mount.

Images were acquired using a laser-scanning confocal microscope (Zeiss, LSM 710), and analysed with ZEN 2012 (Zeiss) and ImageJ 1.48p software. Quantitative analysis of Cux1 expressing cells was done by counting positive cells among all RFP<sup>+</sup> donor cells in the mouse. Cell countings were performed with the Cell Counter plug-in for Image J 1.48p, by careful inspection across serial optical sections (spaced at 1  $\mu\text{m}$ ) of confocal Z-stacks acquired with a 40 $\times$  objective (NA 1.1). Results are represented as mean  $\pm$  s.e.m. calculated between different mice. Image processing was performed with ImageJ and Adobe Photoshop/Illustrator (Adobe Systems) for preparation of multipanel figures.

**Circuit analysis.** For circuits analysis complete brains were carefully removed from the skull following perfusion (see above). After cutting, brain sections were kept in serial order and stained for GFP and RFP. Subsequently, sections with transplanted cells were selected and further stained for cortical layer markers. Of all sections, those with RFP-labelled axonal fibres and/or GFP-labelled cell somas were scanned using an epifluorescence microscope with a motorized stage (Zeiss, Axio ImagerM2) equipped with a 10 $\times$  objective (NA 0.3). We used automatic scanning and alignment of individual tiles, followed by image stitching to create a high resolution image of the whole section. These images were used to identify brain regions where RFP<sup>+</sup> axons or GFP<sup>+</sup> cells were found, by comparison with the corresponding sections of the Allen Reference Atlas of the adult mouse brain<sup>46</sup> (version 2, 2011; website: © 2015 Allen Institute for Brain Science, Allen Mouse Brain Atlas, available from: <http://mouse.brain-map.org>). Some sections of interest are not available in the reference atlas, namely in the sagittal atlas, which displays 21 sections spaced at 200- $\mu\text{m}$  intervals, and only up to 4.0 mm lateral from bregma. In these cases, the Brain Explorer 2 3D viewer (version 2.3.5, Allen Institute for Brain Science, as referenced before) was used to retrieve the corresponding annotated section to overlap it with the experimental section and identify the anatomical location of the labelled cells or axonal fibres. In sections with unclear cell numbers due to close apposition of two GFP somata or with high densities of GFP cells, scanning of confocal Z-stacks with a 40 $\times$  objective (numerical aperture (NA) 1.1) was carried out, and quantification was performed by careful inspection through serial optical sections spaced at 1- $\mu\text{m}$  intervals. In sections containing transplanted cells, four categories were considered for counting: GFP-only cells with neuronal morphology, GFP-only cells with glial morphology, GFP/RFP (or mCherry) cells with neuronal morphology, GFP/RFP (or mCherry) cells with glial morphology. Connectivity ratio was calculated by dividing the number of monosynaptically connected neurons (GFP-only cells with neuronal morphology) in a given region by the number of starter neurons (GFP/RFP or mCherry with neuronal morphology) amongst the transplanted neurons in V1. Results are represented as mean  $\pm$  s.e.m. calculated between different mice.

To assess the efferent projections of transplanted neurons, all brain slices were inspected under an epifluorescence microscope (Zeiss, Axio ImagerM2) ( $n = 3$  mice,  $\sim 134$  slices per brain) using a 20 $\times$  objective (NA 0.8), and the total number of RFP-labelled neurites was counted in each anatomical structure or white matter tract. In a small number of slices per mouse (4–9) quantification of individual fibres could not be performed reliably owing to high background of mCherry staining (as compared to RFP staining) or in densely labelled parts of Vis, cc and ectorhinal cortices, thus the numbers for those structures were expressed as 'greater than'

the number of counted fibres. For these reasons, we express the quantification in categories which we present colour-coded and, which reflect the relative amounts of RFP-labelled neurites in different structures. To determine the normal projectome of upper layer V1 neurons we used data from the Allen Mouse Brain Connectivity Atlas (Allen Brain Atlas Data Portal) that were generated by anterograde tracing using AAV injections in V1. We used data from the Cux2-IRES-Cre mouse line in combination with Cre-dependent AAVs (rAAV2/1-pCAG-flex-eGFP-WPRE-bGH or rAAV2/1-pCAG-flex-synaptophysin-eGFP-WPRE-bGH), to limit the analysis to the projections of upper layer neurons. Projection volumes per structure (sum of detected signal in  $\text{mm}^3$ ) for five individual experiments (501116471, 263780729, 293472335, 483013787 and 501117182; injection volume 0.179–0.384  $\text{mm}^3$ , 98.9–100% in V1) were averaged and values below the threshold of 0.003  $\text{mm}^3$  were not considered. The data were then assembled per anatomical group by summing the projection volumes of the respective structures.

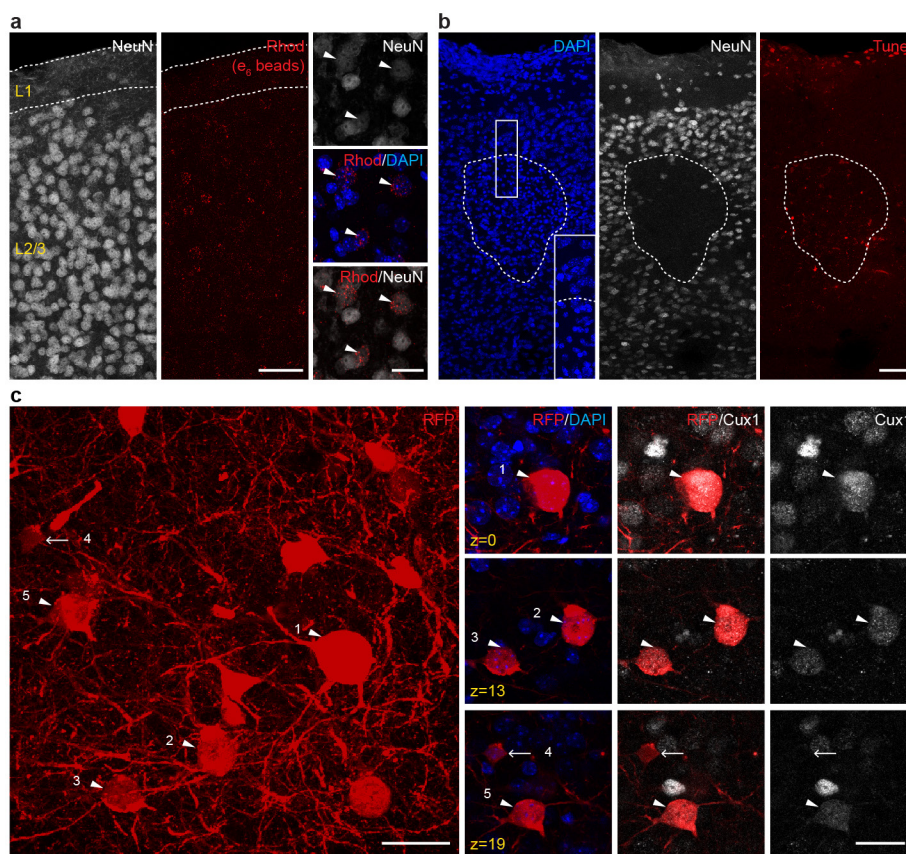
**Analysis of dLGN-V1 topography.** Brain Explorer 2 software was used to retrieve a total of 222 images corresponding to consecutive sagittal sections of an entire hemisphere from the adult mouse brain, where V1 and dLGN volumes were highlighted, using the Allen Mouse Brain Atlas data set. Given the width of a hemisphere (4.2–4.9 mm), each image represents a  $\sim 20.5\ \mu\text{m}$  sagittal slice. Thus, to represent the 60–70  $\mu\text{m}$  thick experimental slices containing labelled neurons, one in every third image was used thereafter. Using Free-D software<sup>47</sup> the V1 and dLGN volumes were reconstructed as 3D surface or wireframe models by manually delineating these structures boundaries as closed contours in the 2D images comprising the stack. For each experimental slice containing starter neurons or pre-synaptic dLGN neurons traced with the RABV, the positions of these neurons were obtained from the Image J Cell Counter images, and scaled down to the whole slice tiled fluorescence images. These were overlapped and aligned with the corresponding anatomical slices of the Allen Brain Atlas data set, and the positions of individual neurons were rendered as points in the Free-D stacks (starter cells in the V1 stack; dLGN neurons in the dLGN stack). To explore the topographic arrangement of dLGN-V1 connections, we plotted the points, starter and dLGN neurons, in a different colour for each mouse ( $n = 8$  mice), in the V1 or dLGN stack, respectively, and rendered them into a 3D representation. For statistical analysis of the topographic relationships we retrieved the  $x$ ,  $y$ ,  $z$  coordinates of each point (cells) as well as each of the points composing the V1 and dLGN contours. Using Matlab R2016a (Mathworks), we calculated the centroid of each group of points (V1, dLGN, V1 starters, dLGN pre-synaptic cells), and plotted the centroids of the V1 or dLGN cell clusters relative to the centroids of each structure, respectively, in order to assess correlations between anterior-posterior, dorso-ventral and medio-lateral dimensions ( $n = 8$  experimental mice, including mice where RABV tracing was performed at 4 wpt (4 mice) or 12 wpt (4 mice), and  $n = 4$  control mice, in which the native circuitry was traced by *in utero* electroporation of the G-TVA plasmid and RABV injection in the adult mouse V1).

**Statistics.** Statistics were performed using PRISM (Graphpad). Appropriate statistical tests were chosen dependent on sample size, data distribution and number of comparisons. Spine and bouton data were analysed with one-way ANOVA and Tukey post-tests for multiple comparisons. Survival curves were analysed pairwise using the Gehan–Breslow–Wilcoxon test, and  $P$ -value thresholds were adjusted for multiple comparisons applying a Bonferroni correction. dLGN-V1 topographic relationships were analysed by linear regression and correlative analysis between each two dimensions in space ( $R^2$  reflects the goodness-of-fit;  $P < 0.05$ ). Functional data were subjected to non-parametric tests using Kruskal–Wallis with Dunn's post-tests. In addition, the tuning of individual neurons across time points, and the average normalized tuning of neurons at 6, 9 and 15 wpt were compared with two-way ANOVA and Tukey post-tests for multiple comparisons. The minimum level of significance was defined as  $P < 0.05$  and all values are reported as mean  $\pm$  s.e.m., if not stated otherwise.

**Data availability.** The data that support the findings of this study are available from the corresponding author upon reasonable request.

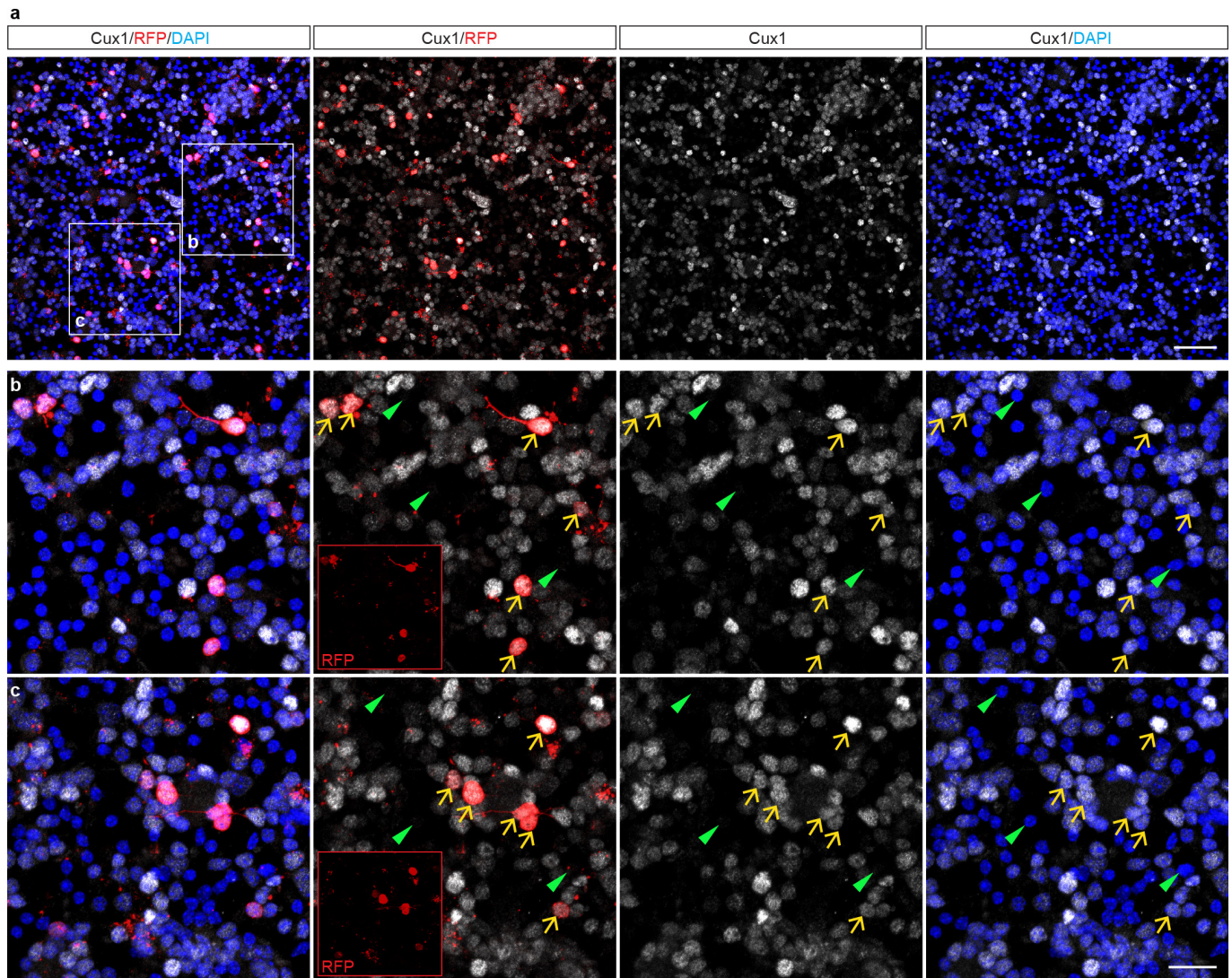
42. Nakamura, T., Colbert, M. C. & Robbins, J. Neural crest cells retain multipotential characteristics in the developing valves and label the cardiac conduction system. *Circ. Res.* **98**, 1547–1554 (2006).
43. Brill, M. S. *et al.* Adult generation of glutamatergic olfactory bulb interneurons. *Nat. Neurosci.* **12**, 1524–1533 (2009).
44. Saito, T. *In vivo* electroporation in the embryonic mouse central nervous system. *Nat. Protocols* **1**, 1552–1558 (2006).
45. Mazurek, M., Kager, M. & Van Hooser, S. D. Robust quantification of orientation selectivity and direction selectivity. *Front. Neural Circuits* **8**, 92 (2014).
46. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
47. Andrey, P. & Maurin, Y. Free-D: an integrated environment for three-dimensional reconstruction from serial sections. *J. Neurosci. Methods* **145**, 233–244 (2005).
48. Rose, T., Goltstein, P. M., Portugues, R. & Griesbeck, O. Putting a finishing touch on GECIs. *Front. Mol. Neurosci.* **7**, 88 (2014).





**Extended Data Figure 1 | Lesion model and identity of transplanted neurons.** **a**, Left, accumulation of Ce6 rhodamine beads (red) around neuronal cell nuclei (NeuN, white) in V1 contralateral to the beads injection. Right, high-magnification panels with DAPI<sup>+</sup> (blue) nuclei, not all of which contain red fluorescent beads, but those that do are NeuN<sup>+</sup> neurons (arrowheads). **b**, TUNEL staining (red) performed 3 days after laser illumination overlaps with loss of NeuN immunoreactivity (white),

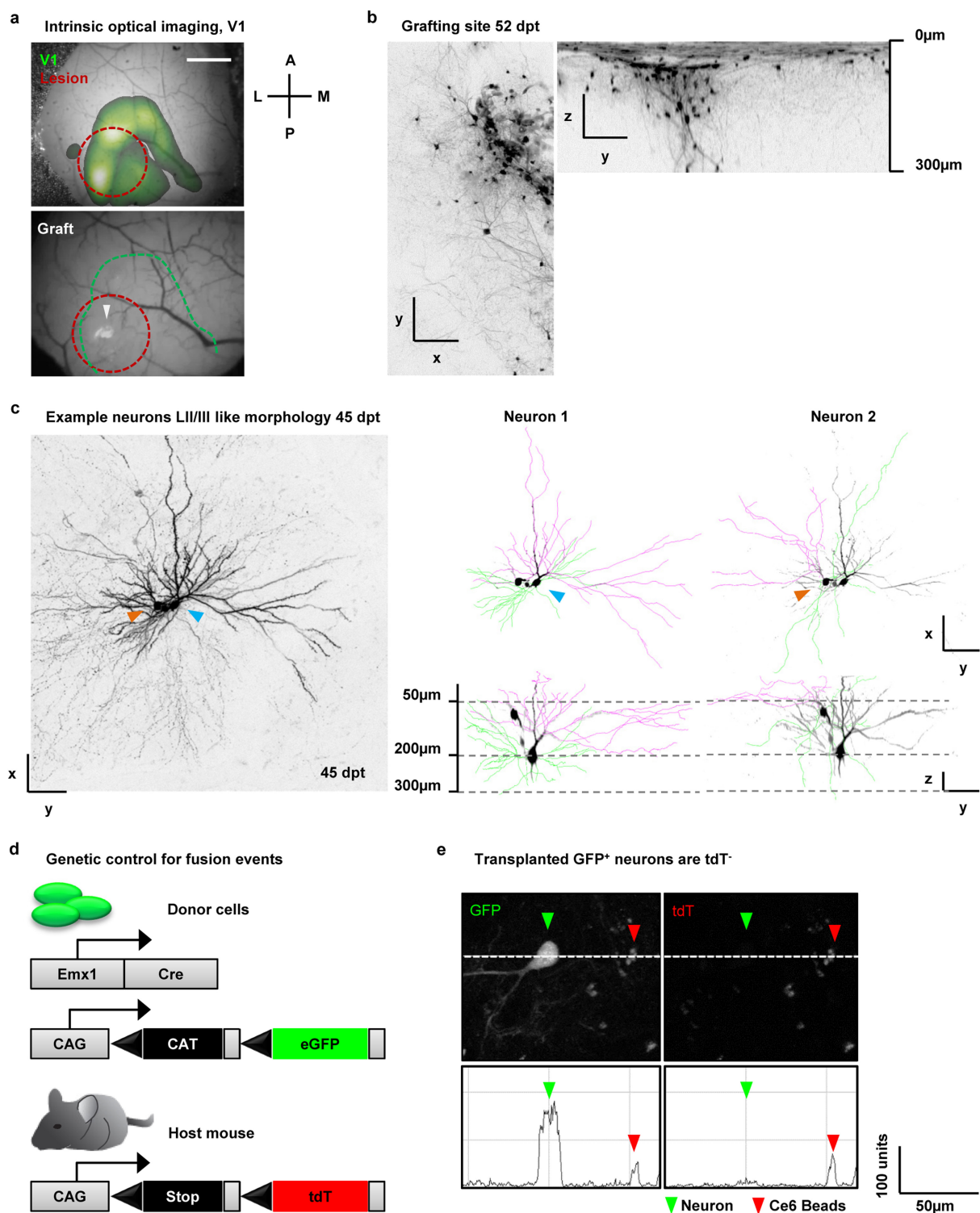
indicating degenerating neurons, and accumulation of condensed apoptotic-like nuclei (DAPI, blue; inset shows high magnification). **c**, Z-stack projection (left) and examples of cells in single optical sections (right) shows that most transplanted neurons express Cux1 (white; arrowheads, Cux1<sup>+</sup> cells; arrow, Cux1<sup>-</sup> cell). Scale bars, 50 μm (**a**, left, **b**), 20 μm (**a**, right), 25 μm (**c**).



**Extended Data Figure 2 | Identity of embryonic neurons before transplantation.** **a**, *In vitro* Cux1 staining (white) on acutely dissociated cells from E18.5 cortex previously labelled by *in utero* electroporation at E14.5 with an RFP-expressing plasmid (red). **b**, **c**, Insets shown in **a**.

Green arrowheads show examples of Cux1<sup>-</sup> cells (nuclear staining in blue, DAPI), which validate the specificity of the immunolabelling; yellow arrows highlight examples of Cux1<sup>+</sup> cells, which are the majority among RFP<sup>+</sup> neurons. Scale bars, 50  $\mu$ m (**a**) and 20  $\mu$ m (**b**, **c**).



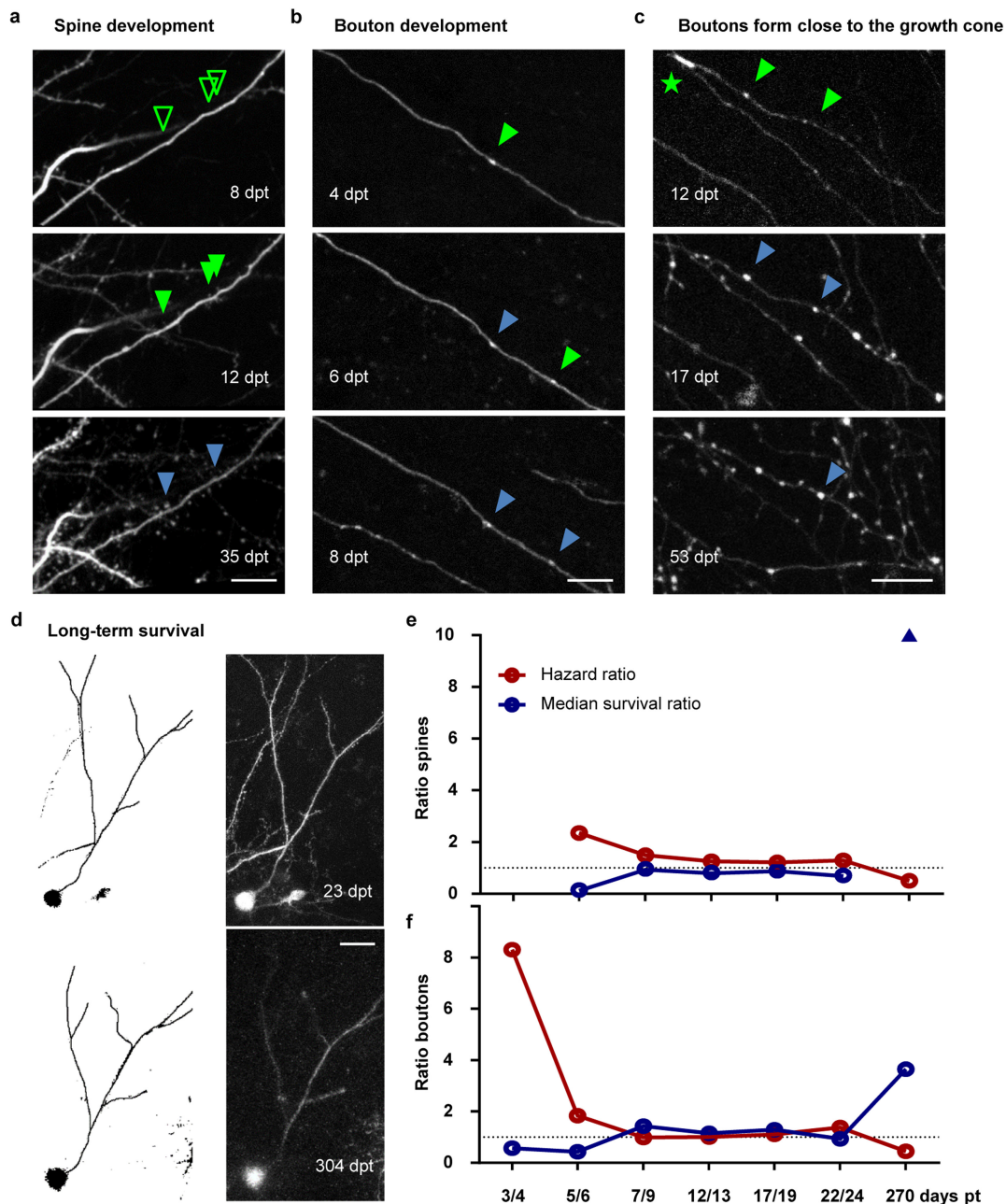


### Extended Data Figure 3 | Intrinsic signal imaging and location of grafting site, neuron reconstruction, and control for fusion events.

**a**, Top, overlay of visual stimulus-evoked intrinsic signal (colour-coded in green) and the blood vessel pattern through a cranial glass window above V1. Red dotted line shows area of laser photoactivation. Bottom, wide-field fluorescence image through the same cranial window (V1, green dotted line). Grafting site (GFP<sup>+</sup>, white arrowhead) in the binocular region of V1. **b**, *In vivo* two-photon z-stack projection (inverted) of a grafting site at 52 dpt, top and side view. **c**, Left, *in vivo* two-photon z-stack projection of a grafting site at 45 dpt (inverted; same as in Fig. 1c). Right, reconstructions (skeleton) of example neurons present in the grafting

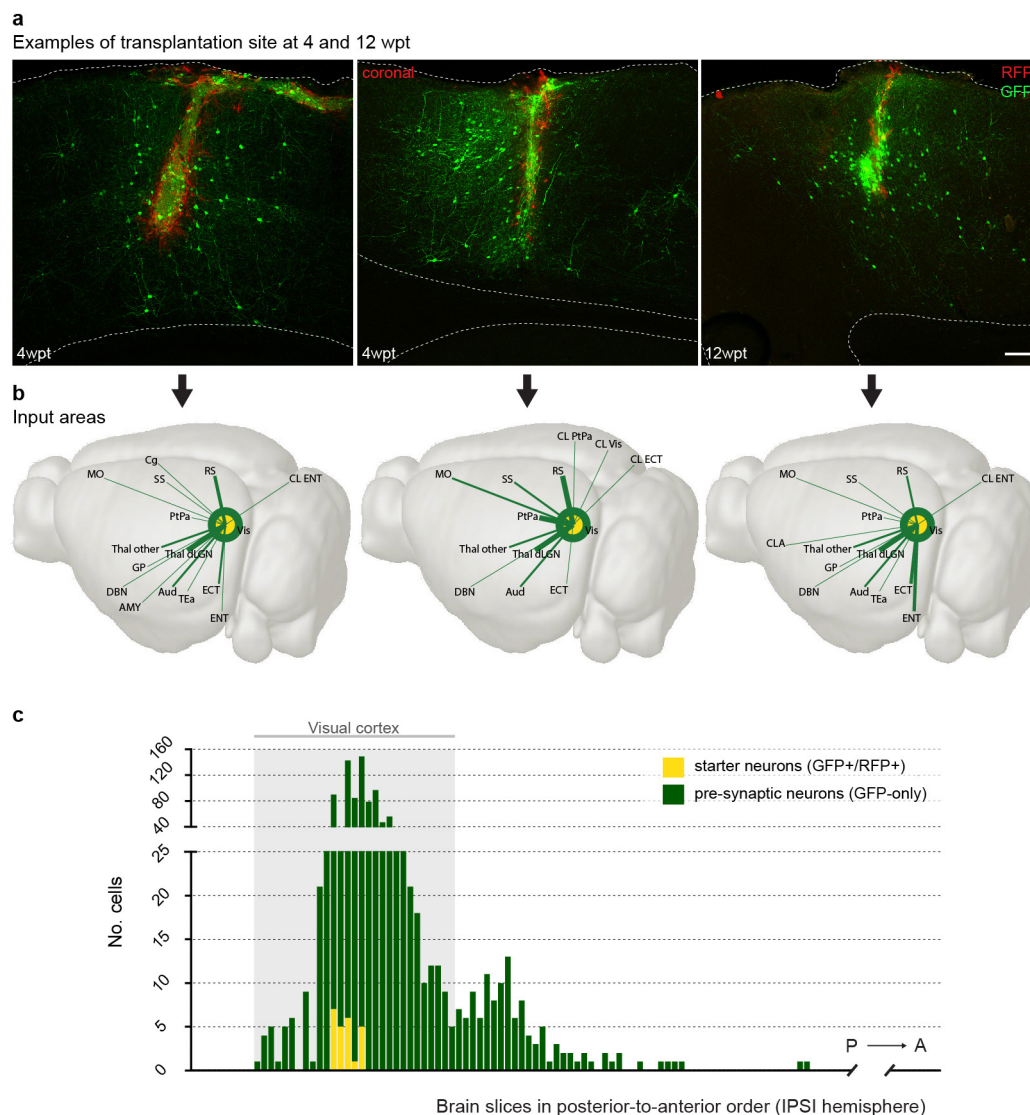
site depicted on the left reveals typical layer 2/3 like morphology. Apical dendrites (magenta) branch from one prominent main dendritic trunk and extend to the surface. Basal dendrites (green) extend from the cell body and reach 300 μm below the pial surface. Top row, top view; bottom row, side view. **d**, Genetic strategy to control for fusion events. Emx1-Cre-driven GFP<sup>+</sup> donor cells were transplanted into tdTomato reporter mice ( $n = 10$ ). **e**, Absence of tdTomato fluorescence in GFP<sup>+</sup> grafted neurons *in vivo*. Line plot (along the white dotted line) across a GFP<sup>+</sup> example neuron (green arrowhead) shows that the neuron is tdT negative. Rhodamine beads (red arrowheads) are equally detected in both channels. Units, 8-bit greyscale. Scale bars, 100 μm (**a**, **b**), and 50 μm (**c**, **e**).





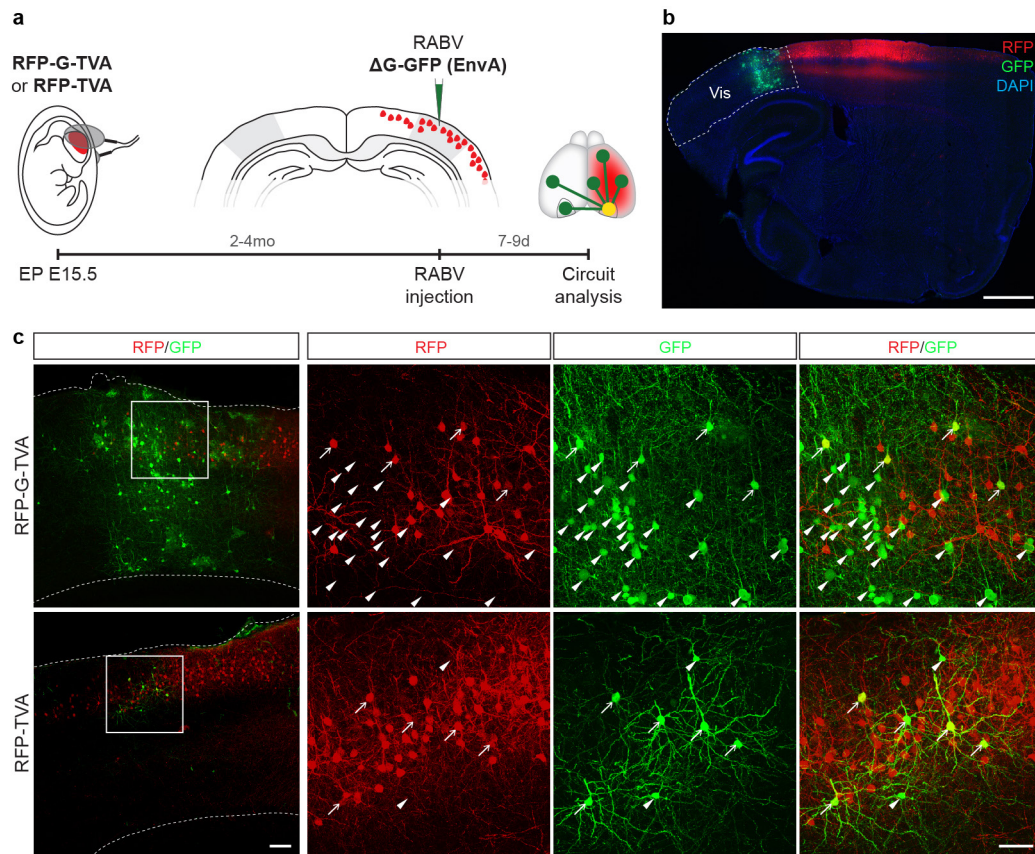
**Extended Data Figure 4 | Modes of formation of spines and boutons, and long-term survival of grafted neurons.** **a**, Dendrites extend and arborize considerably before the first spines form on bare dendrites. Example of a naked dendrite bare from 4 dpt (not shown) to 8 dpt that forms the first spines 12 dpt (empty green arrowheads indicate the location on the dendrite and filled green arrowheads indicate the newly formed spines). Two spines remain stable until 35 dpt (blue arrowheads). **b**, Bouton formation precedes spine formation. Example axon at 4 dpt; arrowheads (green) indicate new boutons that remain stable over subsequent time points (blue arrowheads). **c**, Boutons are able to form within a few micrometres of the axonal growth cone (green star). Individual boutons that have formed in the vicinity of a growth cone are able to survive for days and weeks (blue arrowheads). **d**, Grafted neurons that survived the early phase of integration (>12 dpt) remained stable

until the end of the experiment (here 10 months) and probably for the rest of the animals' life. **e**, **f**, Comparison of early- and late-formed spines and boutons with structures formed at 4–9 wpt. Median survival ratios indicate the relative survival. Hazard ratios indicate the relative chance for structures to be lost. **e**, Early formed dendritic spines (<12 dpt) have a 1.5–2.4 times higher chance of being eliminated compared to spines newly formed at 4–9 wpt. Hazard ratios remain increased (>1.2) up to 24 dpt. Spines formed at 9 mpt, however, have a very high chance of survival. Although the median survival at 4–9 wpt is 28 days, more than half of the spines formed at 9 mpt survive at least for 51 days (blue triangle, arbitrary value as survival is >50%). **f**, Early formed axonal boutons (<7 dpt) are 2–8 times more likely to be eliminated compared to boutons newly formed at 4–9 wpt. Boutons formed at 9 mpt, have a 4 times higher chance of survival. Scale bars, 10  $\mu$ m (**a–c**) and 20  $\mu$ m (**d**).



**Extended Data Figure 5 | Local and brain-wide monosynaptic input to transplanted neurons.** **a**, Examples of the transplantation site in three different animals selected to demonstrate that RFP<sup>+</sup> transplanted neurons (left to right: many to few) are consistently surrounded by a large number of GFP-only labelled connecting neurons from the host. **b**, 3D diagram of the brain-wide monosynaptic input connectome for each of the examples in **a**. Shown are the location of starter neurons in V1 (yellow) and the innervating neurons (green), either local (green circle) or distant (green lines, thickness of the line represents the connectivity ratio for a given area and respects the ranges displayed in Fig. 4f). Note the strong input from thalamic nuclei, in particular the dLGN, in all

individual cases. Importantly, the number of areas and of input neurons in a given area correlates with the number of starter neurons in V1 (see Extended Data Fig. 7). Schematics of brains taken from Brain Explorer, Allen Institute for Brain Science. **c**, Distribution of neurons providing synaptic input to transplanted neurons, example traced at 4 wpt. Number of GFP-only (green) and GFP/RFP double-labelled (yellow) cells throughout the transplanted hemisphere. Each bar corresponds to one coronal section, from posterior (left) to anterior (right) coordinates of the mouse brain. Sections including the visual cortex are highlighted in grey. Note the overrepresentation of local connections compared to long-range projections.



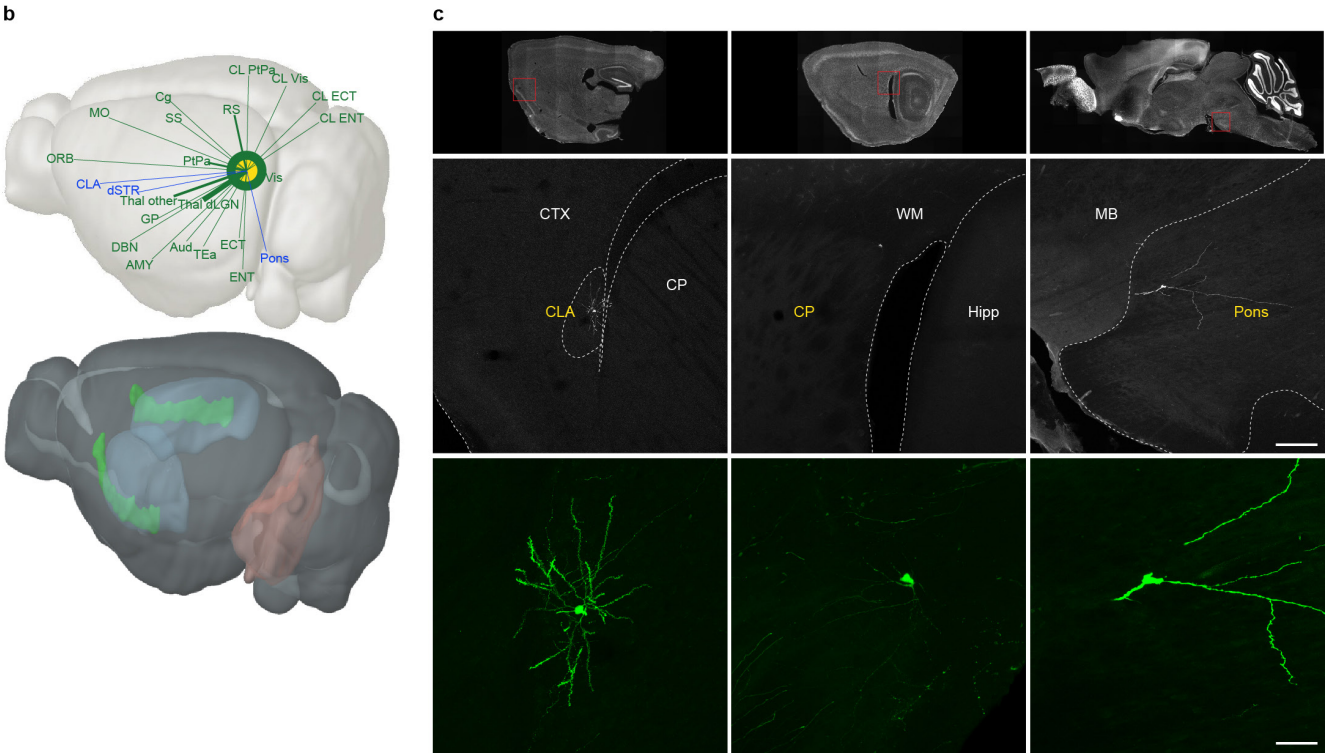
**Extended Data Figure 6 | Normal circuitry of upper layer neurons in V1.** **a**, Schematic depicting experimental procedure. At E15.5, L2/3 neuronal progenitors were *in utero* electroporated to express RFP-G-TVA or RFP-TVA. Electroporation was targeted to the somatosensory cortex to achieve a small number of transduced cells in the nearby visual cortex. Subsequently, in adult mice,  $\Delta$ G-GFP (EnvA) RABV was injected into V1 of the *in utero* electroporated hemisphere. **b**, Sagittal section stained with DAPI (blue) shows RFP-G-TVA electroporated neurons (red) restricted to the upper layers. Axon collaterals of these neurons cause dimmer red

labelling in layer 5. In green, pre-synaptic neurons in and around the area of RABV injection and transduction of few RFP-G-TVA (red) cells. **c**, Higher power micrographs (left) showing RABV-traced cells (green) in V1 with RFP-G-TVA-expressing (top) or RFP-TVA-expressing (bottom) upper layer neurons (red). Right, high magnification shows that a modest number of starter cells (arrows; GFP/RFP) connects robustly with local neurons, while, if primarily infected cells lack G, only occasional and local GFP-only cells are observed (arrowheads). Scale bars, 1 mm (**b**), 100  $\mu$ m (**c**, left) and 50  $\mu$ m (**c**, right).



a

Connectivity	4 wpt				12 wpt				Endogenous			
	Mean	SEM	n (of 6)	Min # starters	Mean	SEM	n (of 6)	Min # starters	Mean	SEM	n (of 3)	Min # starters
<i>IPSI cortical</i>												
Vis	21.502	5.072	6	1	23.089	7.392	6	1	22.528	4.305	3	4
RS	0.485	0.242	4	16	0.442	0.243	3	10	0.304	0.187	2	31
PtPa	0.263	0.248	3	24	0.030	0.019	2	13	0.512	0.250	3	4
Aud	0.087	0.054	3	24	0.076	0.063	2	13	0.052	0.028	2	31
ECT	0.084	0.066	3	24	0.206	0.129	3	10	0.030	0.017	2	31
SS	0.070	0.061	3	24	0.041	0.019	3	10	0.296	0.256	2	31
MO	0.052	0.035	3	24	0.035	0.018	3	10	0.178	0.139	2	31
TEa	0.025	0.020	2	32	0.037	0.026	2	13	0.110	0.070	3	4
ENT	0.013	0.010	2	32	0.151	0.125	3	10	0.022	0.011	2	31
Cg	0.010	0.010	1	32	0.000	0.000	0	—	0.011	0.011	1	31
Orb	0.004	0.004	1	80	0.017	0.017	1	10	0.004	0.004	1	86
<i>IPSI subcortical</i>												
Thal dLGN	1.473	0.574	5	1	2.254	0.763	4	1	1.008	0.159	3	4
Thal other	0.283	0.111	4	3	0.354	0.238	3	10	0.483	0.421	2	31
DBN	0.034	0.017	3	24	0.019	0.013	2	13	0.025	0.020	2	31
GP	0.005	0.005	1	32	0.026	0.026	1	13	0.022	0.022	1	31
AMY	0.005	0.005	1	32	0.000	0.000	0	—	0.000	0.000	0	—
CLA	0.000	0.000	0	—	0.031	0.025	2	13	0.000	0.000	0	—
Pons	0.000	0.000	0	—	0.006	0.006	1	29	0.011	0.011	1	31
CP	0.000	0.000	0	—	0.006	0.006	1	29	0.000	0.000	0	0
MB	0.000	0.000	0	—	0.000	0.000	0	—	0.011	0.011	1	31
<i>CL cortical</i>												
Vis	0.028	0.028	1	24	0.022	0.017	2	10	0.332	0.199	3	4
PtPa	0.014	0.014	1	24	0.000	0.000	0	—	0.011	0.011	1	31
ECT	0.014	0.014	1	24	0.000	0.000	0	—	0.011	0.011	1	31
ENT	0.005	0.005	1	32	0.013	0.013	1	13	0.000	0.000	0	—
SS	0.000	0.000	0	—	0.000	0.000	0	—	0.058	0.052	2	31
MO	0.000	0.000	0	—	0.000	0.000	0	—	0.022	0.022	1	31
Aud	0.000	0.000	0	—	0.000	0.000	0	—	0.015	0.009	2	31
Orb	0.000	0.000	0	—	0.000	0.000	0	—	0.011	0.011	1	31
Tea	0.000	0.000	0	—	0.000	0.000	0	—	0.011	0.011	1	31

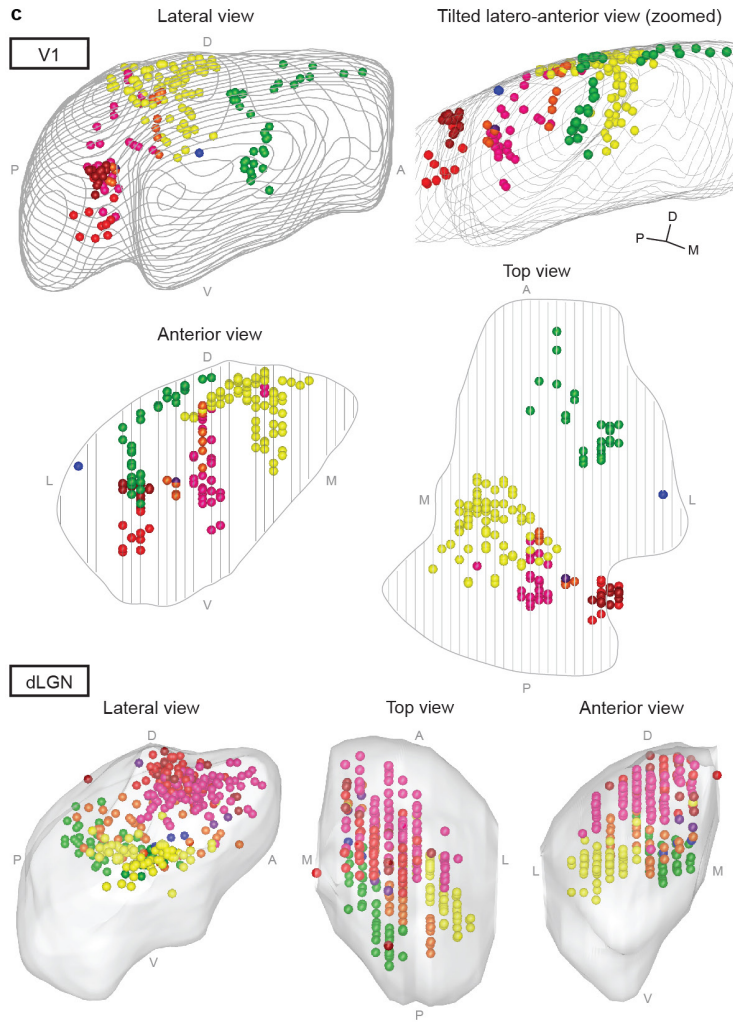
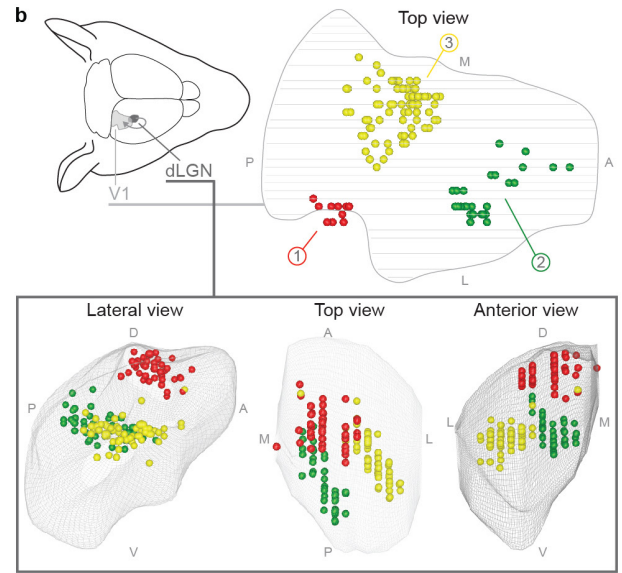
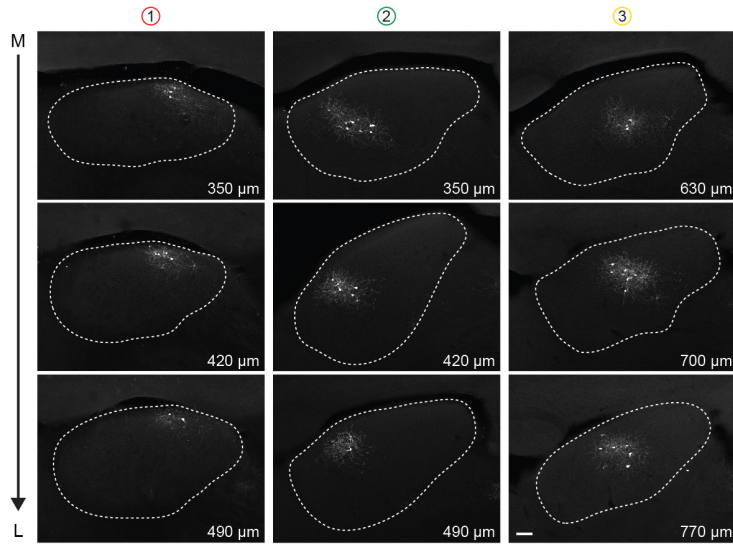
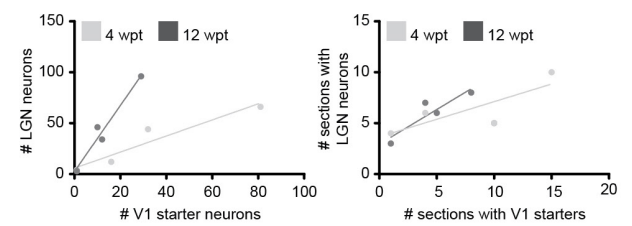
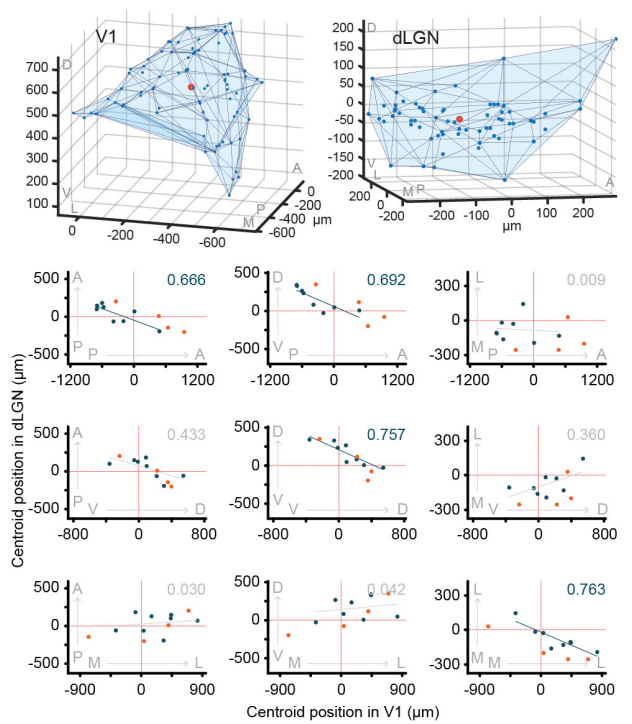


Extended Data Figure 7 | See next page for caption.

### Extended Data Figure 7 | Quantitative analysis of the monosynaptic synaptic tracing of regenerated and endogenous neuronal circuits.

**a**, Connectivity ratio for each anatomical area, obtained from RABV tracing at 4 and 12 wpt, or in control mice, and calculated as described in Methods. The number of transplants/mice with synaptic input from a given area is specified (from a total of 6 mice per group, or 3 control mice in the endogenous circuit tracing). The data indicate that some areas project few axons to V1, and thus a higher number of starters results in increased probability of synapse formation, necessary to unveil these connections; while other areas project massively to V1 and are therefore traced even from only one starter neuron in V1 (Vis, dLGN). The number of starter neurons varied across mice owing to variability of the number of transplanted cells and efficacy of RABV injection (number of starter neurons in each of the 6 mice at 4 wpt: 1, 3, 16, 24, 32 and 80, and in each of the 6 mice at 12 wpt: 1, 1, 3, 10, 13 and 29). Blue shading indicates connections formed only late (12 wpt), grey shading marks the ones that did not form. **b**, Top, 3D diagram of the brain-wide input connections

at 4 wpt (green,  $n = 6$ ) and additional connections revealed at 12 wpt (blue,  $n = 6$ ; regions shown in the bottom diagram, rendered in Brain Explorer 2, Allen Institute for Brain Science.). Shown are the location of starter neurons in V1 (yellow) and the innervating neurons (green/blue), either local (green circle) or distant (green/blue lines; thickness of the line represents the connectivity ratio for a given area and respects the ranges displayed in Fig. 4f). Note the pronounced input from visual cortex and dLGN. **c**, RABV-traced inputs to transplanted neurons observed exclusively at 12 wpt, all known to project to V1 (for dorsal striatum see also online Allen Mouse Brain Connectivity Atlas, experiment 112458831, section 90; experiment 100142580, section 90; experiment 112307754, section 82). Sagittal brain sections with nuclear staining (DAPI) indicate the field magnified below (red square). Yellow text corresponds to regions where GFP-only neurons were observed, and white text denotes nearby regions for anatomical reference. Individual cells are shown at high magnification below. For abbreviations see Extended Data Table 1. Scale bars, 200  $\mu\text{m}$  (**c**) and 50  $\mu\text{m}$  (insets).

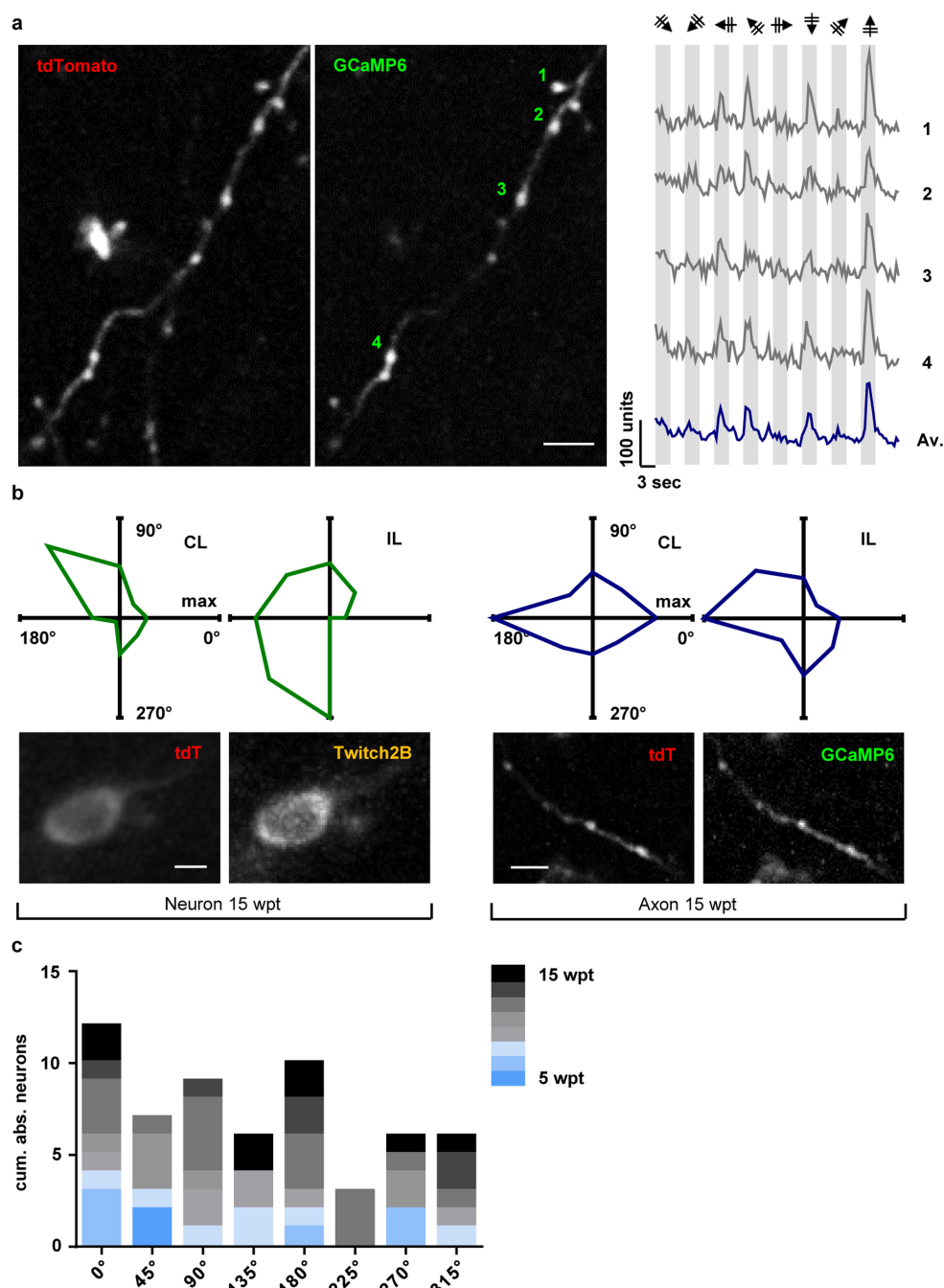
**a** dLGN topography: examples from distinct transplantation areas within V1**d** V1 (graft)-dLGN labeling extent**e** V1 (graft)-dLGN topographic relationships

Extended Data Figure 8 | See next page for caption.



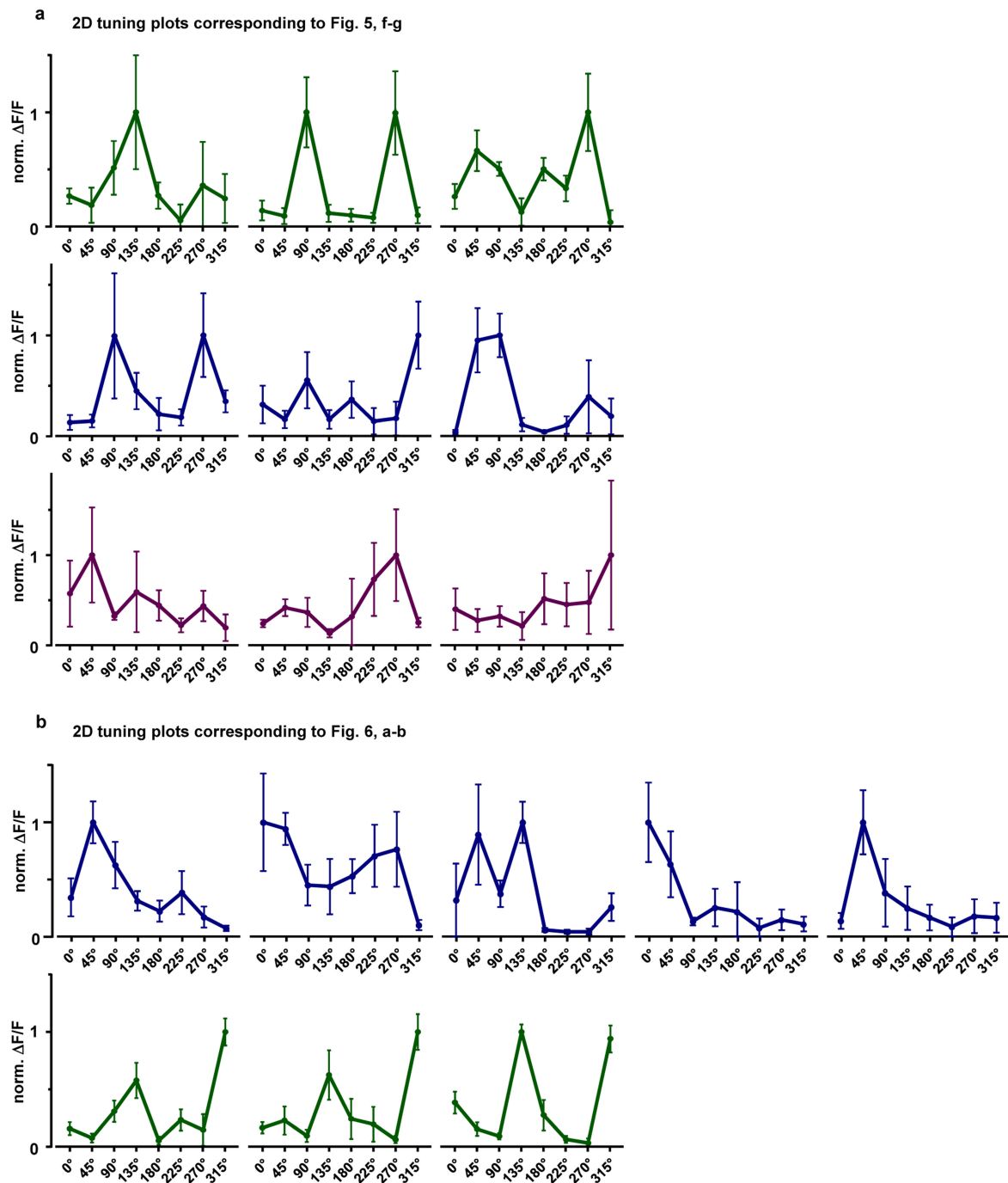
**Extended Data Figure 8 | Transplanted neurons in V1 receive topographically organized inputs from the dLGN.** **a, b,** Examples of dLGN-containing clusters of cells innervating transplants at distinct locations in V1 (mouse 1, 2 and 3), revealed by RABV retrograde tracing. **a,** Each column shows three consecutive sections from medial to lateral (M–L); indicated is the distance to the medial boundary of the dLGN. Dashed lines indicate the outline of the dLGN, dorsal is up, anterior is to the right. **b,** 3D reconstruction of V1 starter cells and their pre-synaptic dLGN cells from mice 1, 2 and 3, plotted into the same rendered V1 (top view) and dLGN (lateral, top and anterior views), and shown in a different colour for each mouse (each point represents one cell). All three clusters receive inputs from topographically corresponding parts of the dLGN. **c,** 3D reconstruction comprising all mice ( $n = 8$ ) analysed shows that pre-synaptic cells in the dLGN always cluster, that cluster size in V1 and dLGN correlates, and the progressive nature of the topography (see intermediate clusters, for example, pink and orange in ML–ML correlation). Tilted latero-anterior view of V1 demonstrates the similar depth of each transplant in V1, reaching down to approximately the middle of the cortical thickness. **d,** Correlation between the extent of labelling in V1 (starter neurons) and the respective dLGN (pre-synaptic neurons),

by quantification either of the total number of cells (left) or of slices containing cells as a measure for cells' dispersion in the cluster (right), at 4 (light grey) and 12 wpt (dark grey). **e,** Top, example of the 3D boundary encompassing a cluster of transplanted cells in V1 and the corresponding pre-synaptic cell cluster in the dLGN. Blue dots represent individual cells, red circle indicates the centroid of each cluster. Bottom, centroid coordinates of V1 and dLGN cell clusters (blue dots) are plotted relative to the centroid of the V1/dLGN rendered volumes ( $x, y, z = 0, 0, 0$ ; intersection of the red lines), for each two dimensions in space (A–P, antero-posterior; M–L, medio-lateral; D–V, dorso-ventral). Orange dots show results from analysis of the native circuitry. Regression line and coefficient of determination ( $R^2$ ; values at top right corner) for the experimental group are indicated for each plot; blue lines/values indicate slopes significantly non-zero, that is, data close to the fitted linear regression ( $P < 0.05$ ); grey lines/values indicate non-significant correlations. Regarding dLGN-to-V1 projections, the dimensions AP–AP, DV–AP and ML–ML inversely correlate, as is known and confirmed in our control data (orange dots). Also, a DV–DV correlation is revealed, although it probably results from overrepresented DV coordinates of cells in V1 due to the V1 curvature. Scale bar, 100  $\mu\text{m}$ .



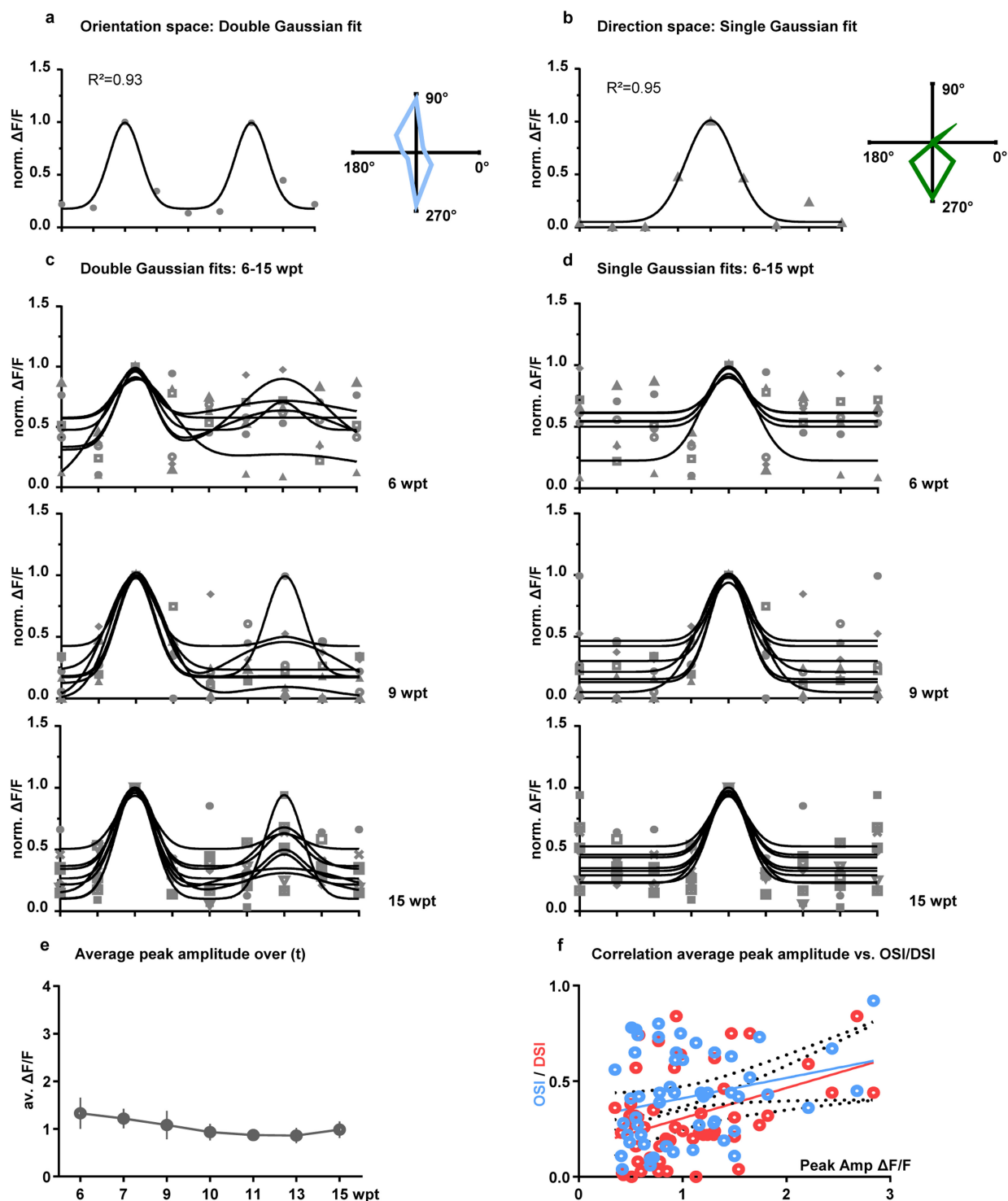
**Extended Data Figure 9 | Responses to visual stimulation: boutons on the same axon, binocular responses, distribution of preferred directions.** **a**, Left, axon of a transplanted neuron (tdTomato<sup>+</sup>) expressing GCaMP6, single optical plane, maximum projection of all frames of one stimulation sequence (see Methods). Right, individual (grey) and average (blue) responses (arbitrary units) of 4 boutons (indicated on the left) to visual stimulation with gratings moving in 8 directions (grey bars, direction indicated on top). Note highly similar responses of all

boutons. **b**, Individual transplanted neurons respond to ipsilateral (IL) and contralateral (CL) eye stimulation. Top, polar plots; bottom, single plane, maximum projection of all frames of one stimulation sequence. **c**, Transplanted neurons are tuned to all directions, with a slight overrepresentation of cardinal directions. Cumulative absolute number of grafted neurons across imaging time points, sorted according to their respective preferred direction. Scale bars, 5  $\mu$ m.



**Extended Data Figure 10 | Normalized average 2D tuning recorded from example neurons, axons and spines. a, b,** Normalized average tuning plots  $\pm$  s.e.m. for the examples presented in Fig. 5f, g (a) and in Fig. 6a, b (b). Same arrangements and time points as in Figs 5 and 6.





**Extended Data Figure 11 | Orientation and direction selectivity assessed with Gaussian fits.** **a**, Example cell displaying strong orientation selectivity (right, polar plot), fitted with a double Gaussian (DG) function. **b**, Example cell displaying strong direction selectivity (right, polar plot), fitted with a single Gaussian (SG) function. **c**, Individual DG fits of 17 cells in total ( $n = 4$  mice) at 6, 9 and 15 wpt. **d**, Individual SG fits of 17 cells in total ( $n = 4$  mice) at 6, 9 and 15 wpt. **e**, **f**, Peak amplitudes are constant and

do not correlate with specificity. **e**, Consistent average peak amplitudes across imaging time points.  $P = 0.6932$  (not significant), Kruskal–Wallis test. **f**, Orientation and direction selectivity indices do not correlate with average peak amplitudes (OSI,  $R^2 = 0.0692$ ; DSI,  $R^2 = 0.1495$ ; dashed lines, 95% confidence interval). Reported<sup>48</sup> nonlinearity effects of calcium indicators would predict higher selectivity indices at lower average peak amplitudes.

**Extended Data Table 1 | Abbreviation list of anatomical areas in the adult mouse brain**

Abbreviation	Anatomical area
ACB	Nucleus accumbens
act	Anterior commissure, temporal limb
AI	Agranular insular area
alv	Alveus
AMY	Amygdala
Aud	Auditory cortex
BLA	Basolateral amygdalar nucleus
BMA	Basomedial amygdalar nucleus
cc	Corpus callosum
CEA	Central amygdalar nucleus
Cg	Cingulate cortex
cing	Cingulum bundle
CLA	Clastrum
CP	Caudoputamen
DBN	Diagonal band nucleus
DG	Dentate gyrus
ECT	Ectorhinal cortex
ENT	Entorhinal cortex
EP	Endopiriform nucleus
FS	Fundus of striatum
GP	Globus pallidus
GU	Gustatory areas
Hipp	Hippocampus
HPF	Hippocampal formation
IA	Intercalated amygdalar nucleus
LA	Lateral amygdalar nucleus
LV	Lateral ventricle
MB	Midbrain
MO	Motor cortex
OB	Olfactory bulb
opt	Optic tract
ORB	Orbital cortex
PALd	Dorsal pallidum
PARA	Parasubiculum
PERI	Perirhinal cortex
PIR	Piriform area
Pons	Pons
POST	Postsubiculum
PRE	Presubiculum
PtPa	Posterior parietal association areas
RS	Retrosplenial cortex
sAMY	Striatum-like amygdalar nuclei
SCs	Superior colliculus
SS	Somatosensory cortex
STRd	Dorsal striatum
STRv	Ventral striatum
SUB	Subiculum
TEa	Temporal association areas
Thal dLGN	Thalamic dorsal lateral geniculate nucleus
Thal other	Other thalamic nucleus
TR	Postpiriform transition area
Vis	Visual cortex
VISC	Visceral area
V1	Primary visual cortex
WM	White matter

# Balancing selection shapes density-dependent foraging behaviour

Joshua S. Greene<sup>1</sup>, Maximillian Brown<sup>1</sup>, May Dobosiewicz<sup>1</sup>, Itzel G. Ishida<sup>1</sup>, Evan Z. Macosko<sup>1</sup>, Xinxing Zhang<sup>2</sup>, Rebecca A. Butcher<sup>2</sup>, Devin J. Cline<sup>3</sup>, Patrick T. McGrath<sup>3</sup>§ & Cornelia I. Bargmann<sup>1</sup>§

**The optimal foraging strategy in a given environment depends on the number of competing individuals and their behavioural strategies. Little is known about the genes and neural circuits that integrate social information into foraging decisions. Here we show that ascaroside pheromones, small glycolipids that signal population density, suppress exploratory foraging in *Caenorhabditis elegans*, and that heritable variation in this behaviour generates alternative foraging strategies. We find that natural *C. elegans* isolates differ in their sensitivity to the potent ascaroside icas#9 (IC-asc-C5). A quantitative trait locus (QTL) regulating icas#9 sensitivity includes *srx-43*, a G-protein-coupled icas#9 receptor that acts in the ASI class of sensory neurons to suppress exploration. Two ancient haplotypes associated with this QTL confer competitive growth advantages that depend on ascaroside secretion, its detection by *srx-43* and the distribution of food. These results suggest that balancing selection at the *srx-43* locus generates alternative density-dependent behaviours, fulfilling a prediction of foraging game theory.**

The success of a particular foraging strategy varies according to the behaviour of competitors. Balancing selection can therefore favour the co-existence of multiple strategies within a species<sup>1,2</sup>. The pioneering example of strategic competition is the natural genetic variation seen at the *foraging* (*for*) gene in *Drosophila melanogaster* larvae<sup>3,4</sup>. Two *for* alleles for active (rover) or sedentary (sitter) behaviour are maintained in a population because of frequency-dependent balancing selection against larvae with the more common foraging strategy<sup>5</sup>. This example, and others like it, suggests that animals could benefit from detecting and responding to competitors in real time and modifying foraging behaviour accordingly. However, little is known about the genes and neural circuits that incorporate information about conspecifics into foraging strategies.

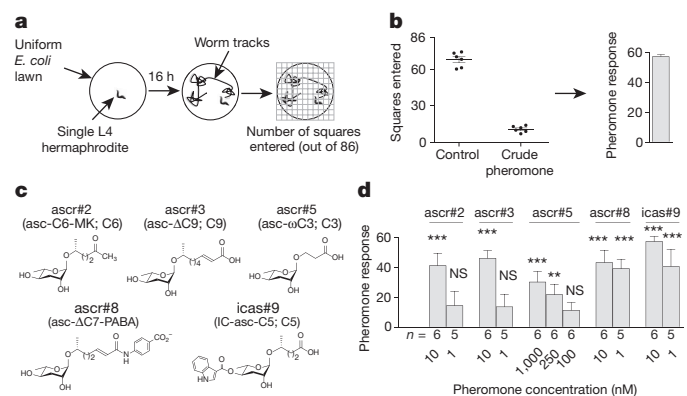
Owing to the well-characterized nature of its foraging circuits and intraspecific pheromone system, the nematode *C. elegans* provides an opportunity to address this question. The *C. elegans* strategy for foraging bacterial food spontaneously alternates between an exploratory behaviour (roaming) and a less active behaviour (dwelling), with each persisting for several minutes per episode<sup>6,7</sup>. Transitions between roaming and dwelling are regulated by distributed neuromodulatory systems that link internal cues, such as nutritional status, to locomotion circuits<sup>7,8</sup>. *C. elegans* senses population density using a family of secreted pheromones called ascarosides, molecules that control the developmental decision to enter the starvation-resistant dauer larva stage<sup>9</sup> and regulate behaviours such as aggregation and male attraction to hermaphrodites<sup>10,11</sup>. We show here that physiological levels of certain ascarosides also regulate foraging by suppressing roaming behaviours. By characterizing differences in pheromone sensitivity in natural *C. elegans* isolates, we identify a pheromone receptor that shapes alternative foraging strategies and affects fitness depending on the structure of the food environment.

## Pheromones regulate foraging behaviour

The effects of ascaroside pheromones on *C. elegans* foraging were examined by quantifying long-term exploration of a bacterial lawn by

individual wild-type N2 strain animals (Fig. 1a). To mimic the effects of high population density on these isolated animals, we conducted the assay in the presence of natural pheromone extracts. The pheromones strongly suppressed exploration (Fig. 1b), as did several pure synthetic ascarosides at concentrations at or below those that induced dauer larva development (Fig. 1c, d). However, *ascr#5*, a potent regulator of dauer development, only weakly suppressed exploration (Fig. 1d). Thus a subset of ascarosides regulates foraging behaviour at biologically relevant concentrations.

The exploration assay is an indirect measure of the relative time *C. elegans* spends in roaming and dwelling states<sup>7,8</sup>. Quantitative behavioural analysis of video recordings showed that the potent ascarosides icas#9 and *ascr#8* decreased the fraction of time spent

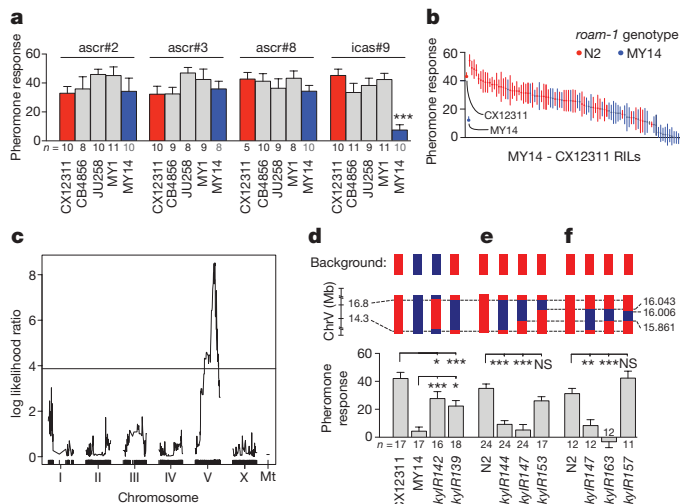


**Figure 1 | Ascaroside pheromones suppress exploratory foraging behaviour.** **a**, Schematic of the exploration assay. **b**, Wild-type N2 response to crude pheromone extract, showing exploration scores and a pheromone-response index, presented as mean  $\pm$  s.e.m. **c**, Structures and names of selected ascarosides. **d**, N2 response to individual ascarosides, presented as mean  $\pm$  s.e.m.. \*\* $P < 0.01$ , \*\*\* $P < 0.001$  by ANOVA with Dunnett's correction; NS, not significant.

<sup>1</sup>Howard Hughes Medical Institute, Lulu and Anthony Wang Laboratory of Neural Circuits and Behavior, The Rockefeller University, New York, New York 10065, USA. <sup>2</sup>Department of Chemistry, University of Florida, Gainesville, Florida 32611, USA. <sup>3</sup>School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, USA.

§These authors jointly supervised this work.





**Figure 2 | Natural genetic variation in pheromone sensitivity.**

**a**, Response of wild-type *C. elegans* strains (x axis) to synthetic ascarosides. \*\*\* $P < 0.001$  by ANOVA with Dunnett's correction. **b**, icas#9 response of 94 CX12311-MY14 RILs and parental strains. **c**, QTL analysis of RILs shown in **b**. Horizontal line denotes  $P < 0.05$  genome-wide significance threshold. Mt, mitochondrial. **d–f**, Mean icas#9 response of NILs used to map the *roam-1* locus, in MY14 and CX12311 (**d**) or N2 (**e**, **f**) backgrounds. \*\* $P < 0.01$ , \*\*\* $P < 0.001$  by ANOVA with Dunnett's correction. All data are presented as mean  $\pm$  s.e.m.

roaming, the duration of roaming states and the speed of locomotion during roaming, but did not affect dwelling states (Extended Data Fig. 1).

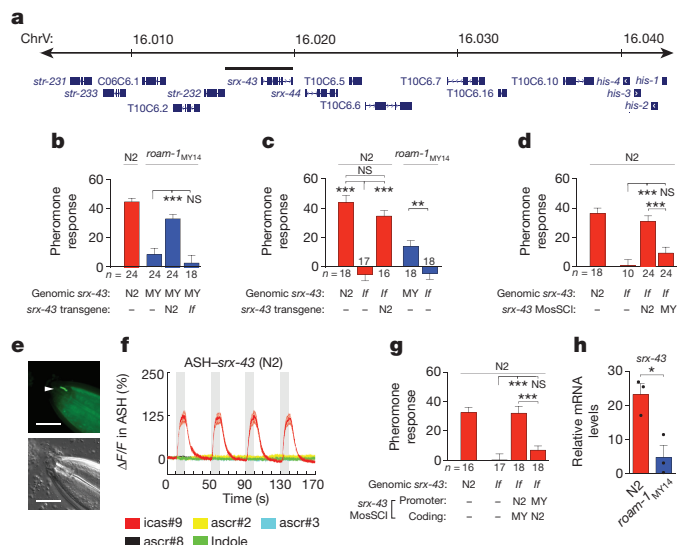
### A QTL for pheromone sensitivity

A variety of genetically diverse wild-type *C. elegans* strains, including the control N2-like strain CX12311, responded to the presence of ascarosides with a suppression of exploration<sup>12</sup> (Fig. 2a). The wild-type German strain MY14 failed to respond to 10 nM icas#9 in the exploration assay, although it responded normally to the presence of ascr#2, ascr#3, and ascr#8 (Fig. 2a and Extended Data Fig. 2). Coupled alterations in the levels of pheromone signalling and detection ability can contribute to reproductive isolation during incipient speciation<sup>13</sup>. However, MY14 and CX12311 were found to produce similar levels of icas#9 and another 16 ascarosides (Extended Data Fig. 3), indicating that the change in the icas#9 response in MY14 was independent of icas#9 production.

To determine the genetic basis of icas#9 insensitivity, 94 recombinant inbred lines (RILs) were generated from intercrosses between MY14 and CX12311 strains. A continuous distribution of icas#9 sensitivity was observed in the exploration behaviour of the RILs (Fig. 2b), suggesting that two or more loci contribute to icas#9 sensitivity. The 94 RILs were genotyped at  $\sim 185$ -kb resolution across the genome by low-coverage whole-genome sequencing<sup>14</sup> (Supplementary Table 1). QTL analysis identified a single significant QTL of genome-wide significance that accounted for 34.9% of the total variance between the RILs, which we term *roam-1* (Fig. 2c). Covariate analysis failed to find additional QTLs that were either additive or interactive with *roam-1* (Extended Data Fig. 4).

The impact of *roam-1* on foraging was confirmed by creating near-isogenic lines (NILs) that had small genetic regions substituted between the strains. NILs in which the 2.5 Mb surrounding *roam-1* were reciprocally exchanged between CX12311 and MY14 were intermediate in their icas#9 sensitivity compared to the parental strains (Fig. 2d). Depending on the direction of the introgression, the *roam-1* locus accounted for 39–46.4% of the genetic variance between the two parental strains.

To simplify further mapping, the *roam-1* region from MY14 (*kyIR139*) was crossed into the N2 laboratory strain. The resulting NIL



**Figure 3 | The *roam-1* QTL includes the icas#9 receptor SRX-43.** **a**, The *roam-1* locus. **b–d**, Sensitivity to icas#9, as affected by high-copy *srx-43* transgenes (**b**), *srx-43(lf)* mutations (**c**) or single-copy integrated *srx-43* transgenes (**d**). **e**, Image showing *srx-43::GFP* translational reporter (top) and Nomarski image (bottom). Arrowhead indicates an ASI sensory cilium. Scale bar, 10  $\mu$ m. **f**, ASH calcium responses in ASH::*srx-43* transgenic strain. **g**, icas#9 sensitivity after exchanging promoters in single-copy *srx-43* transgenes. **h**, Endogenous *srx-43* mRNA levels ( $n = 3$  assays each, indicated). All data are presented as mean  $\pm$  s.e.m. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  by ANOVA with Dunnett's correction or *t*-test.

(*kyIR144*, Fig. 2e) facilitated further mapping that localized *roam-1* to 182 kb (Fig. 2e). High-density mapping of 2,600 F2 progeny of a cross between N2 and the NIL *kyIR147* yielded 12 informative recombinants in this 182-kb region that mapped the *roam-1* QTL to a 37-kb region on chromosome V (Fig. 2f and Methods).

### *roam-1* affects the icas#9 receptor SRX-43

The 37-kb *roam-1* region contains 16 protein-coding genes, including 5 genes that encode predicted G-protein-coupled chemoreceptors in the *srx* or *str* gene families (Fig. 3a). We thought it possible that one or more of the chemoreceptors could be icas#9-receptors, with reduced activity in MY14. We therefore introduced N2-derived sequences that overlapped the chemoreceptor genes into the *roam-1*<sub>MY14</sub> NIL strain *kyIR163*. N2-derived transgenes that covered *srx-43* (Fig. 3a, bar) conferred sensitivity to icas#9 in *roam-1*<sub>MY14</sub> organisms (Fig. 3b), whereas transgenes with nonsense mutations disrupting the coding region of *srx-43* did not (Fig. 3b).

The function of *srx-43* was examined further by characterizing loss-of-function mutations in the endogenous *srx-43* gene (*srx-43(lf)*; Methods). We found that *srx-43(lf)* mutants in both N2 and *roam-1*<sub>MY14</sub> genetic backgrounds were insensitive to icas#9 (Fig. 3c). Sensitivity to icas#9 in the N2 *srx-43(lf)* mutant was restored by the introduction of an N2 *srx-43* transgene (Fig. 3c). These results indicate that *srx-43* is necessary for the icas#9 response in both N2 and MY14 strains and is essential for the behavioural difference between them.

The activity of the N2 and MY14 *srx-43* genes was compared by targeting a single copy of *srx-43* from each strain to a defined locus. This was done using the Mos1 transposase in an *srx-43(lf)* mutant, so that the single-copy transgene was the sole source of *srx-43*. The N2 *srx-43* genomic region fully rescued the icas#9 response, whereas the MY14 region did not (Fig. 3d). The different effects induced by the introduction of different single-copy transgenes indicate that MY14 *srx-43* possesses reduced activity compared to N2 *srx-43*.

Reporter genes with N2 or MY14 *srx-43* sequences driving expression of green fluorescent protein (GFP) were expressed selectively in the ASI sensory neurons (Extended Data Fig. 5a), which promote roaming

behaviour<sup>8,15</sup>. A genomic clone with GFP fused to the C terminus of the SRX-43 protein was enriched in ASI sensory cilia, the site of sensory transduction (Fig. 3e), suggesting that SRX-43 is a chemoreceptor. We investigated the effects of *icas#9* on ASI activity using *in vivo* calcium imaging, but did not observe a response. This negative result is consistent with studies of dauer formation, in which ascarosides regulate gene expression in ASI neurons and not acute calcium levels<sup>16–19</sup>.

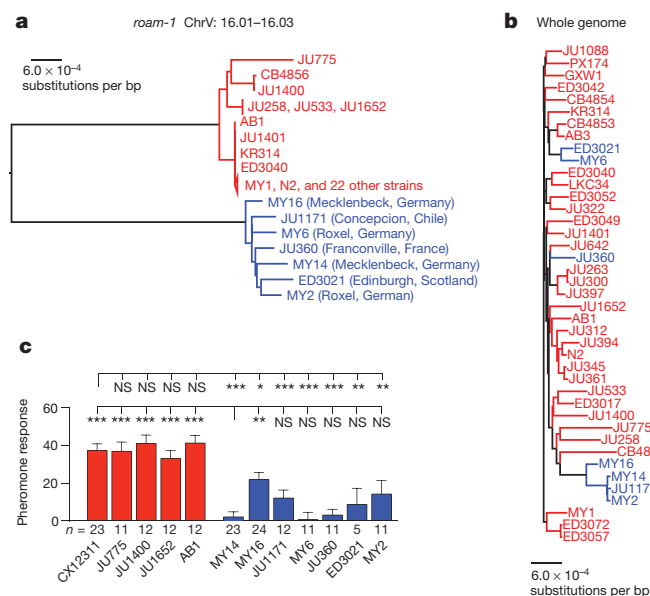
To investigate whether SRX-43 was an *icas#9* receptor, *srx-43* cDNA was expressed in the ASH class of sensory neurons, which are normally insensitive to ascarosides<sup>17</sup> (Extended Data Fig. 5b). Ascaroside-induced calcium flux was then monitored using genetically encoded calcium indicators<sup>17,20</sup>. ASH neurons expressing SRX-43 responded to 10 nM *icas#9* with calcium transients, but this was not observed in response to other ascarosides or indole (Fig. 3f). Although the MY14 strain was largely insensitive to *icas#9* in foraging assays, MY14 SRX-43 also detected *icas#9* when expressed in ASH (Extended Data Fig. 5c).

We observed expression of a MY14-derived *srx-43::GFP* reporter gene in ASI neurons, but this appeared to be weaker than expression of the N2 *srx-43::GFP* reporter (Extended Data Fig. 5a). To investigate which sequences led to the difference in N2 and MY14 *srx-43* activity, we exchanged the *srx-43* promoter and coding regions of both strains and tested these constructs as Mos1-mediated single-copy insertion (MosSCI) *srx-43* transgenes. A transgene with the promoter region of the N2 strain and the coding region of MY14 rescued *icas#9* sensitivity in *srx-43(lf)* mutants but the opposite did not, localizing the difference to the *srx-43* promoter (Fig. 3g). Quantitative measurements of endogenous *srx-43* mRNA levels demonstrated that *srx-43* was expressed at a level fivefold lower in *roam-1<sub>MY14</sub>* than in N2 (Fig. 3h). Therefore, the natural variation in *srx-43* promoter activity between N2 and MY14 affects *srx-43* gene expression and behavioural sensitivity to *icas#9*.

Ascarosides promote dauer larva development in part by suppressing the transcription of *daf-7*, a gene expressed in ASI neurons that encodes a secreted TGF $\beta$ -related peptide<sup>19</sup>. As with animals treated with ascarosides, *daf-7(lf)* mutants exhibit reduced levels of roaming<sup>7</sup>. Expression of a *daf-7::GFP* reporter was significantly reduced by treatment with *icas#9*, indicating that *daf-7* may be a target for *icas#9*; animals bearing a *daf-7* mutation were also less responsive to *icas#9* than controls (Extended Data Fig. 5d, e). Behaviour was only influenced by *icas#9* after several hours of exposure (Extended Data Fig. 5f), a delay that is in agreement with the slow transcriptional regulation of the *daf-7* signalling pathway. Together, these results suggest that *icas#9* acts as a primer pheromone<sup>21,22</sup> that regulates foraging via transcription and endocrine signalling.

### Balancing selection at a foraging QTL

To understand the population genetics of *roam-1*, we examined the genomic sequence of a 20-kb region centred around *srx-43* in 39 additional wild-type *C. elegans* isolates sequenced by the Million Mutation Project<sup>23</sup>. Two discrete, highly divergent haplotypes for the *roam-1* region were found: one resembling N2 that was present in 34 strains and another resembling MY14 that was present in 7 strains with different geographical origins and genetic backgrounds (Fig. 4a, b and Extended Data Fig. 6a). The N2 and MY14 haplotypes differed at 2.64% of all positions over the 20-kb *srx-43* region, a figure 12 times that of the genome-wide average<sup>24</sup>. These data were, however, derived using Illumina sequencing, a method that can fail to align highly divergent sequences. Targeted Sanger sequencing revealed that MY14 and N2 actually differed at 19.7% of all positions in the *srx-43* promoter and coding regions (Extended Data Fig. 6b). A phylogeny constructed for *srx-43* and the most closely related genes in *C. elegans*, *C. briggsae*, and *C. remanei* confirmed that the divergent *srx-43* alleles represent the same gene (Extended Data Fig. 6c). All seven tested MY14-like strains were relatively resistant to *icas#9* compared to N2-like strains (Fig. 4c). These results suggest that naturally occurring resistance to



**Figure 4 | Population genetics of the *roam-1* locus and *icas#9* sensitivity.**

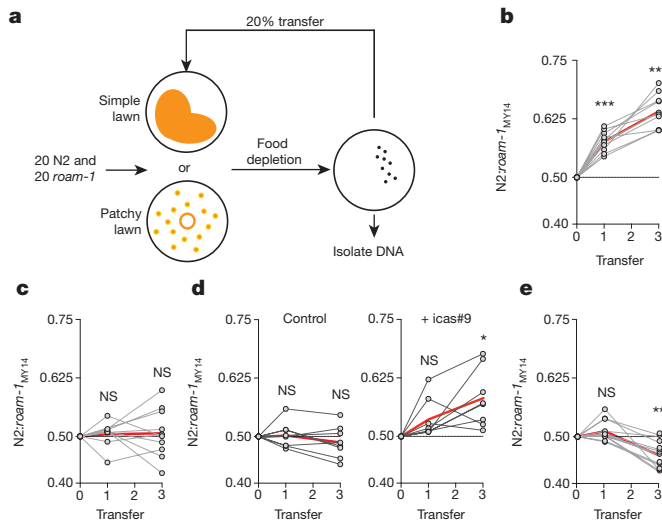
**a**, A dendrogram across 41 natural isolates representing a 20-kb region surrounding *srx-43*. The two major *roam-1* haplotypes from N2 and MY14 strains are indicated in red and blue, respectively. **b**, Whole-genome dendrogram for the same strains as **a**, showing relationships among strains. Red and blue colours follow *roam-1* haplotypes in **a**. **c**, Responses to *icas#9* in natural isolates with the MY14 haplotype (blue) are consistently lower than those with the N2 haplotype (red). Pheromone response shown as mean  $\pm$  s.e.m. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  by ANOVA with Dunnett's correction; NS, not significant.

*icas#9* is associated with a highly divergent *roam-1* haplotype that includes *srx-43*.

The marked allelic divergence of the *roam-1* region, together with the observation that most genes in the interval including *srx-43* have a low  $d_N/d_S$  ratio (Extended Data Table 1), suggests that these alternative haplotypes might be subject to balancing selection. To examine this possibility, we analysed a large database of wild-type strains assembled and sequenced by the Andersen laboratory at Northwestern University (CeNDR; <http://www.elegansvariation.org>). The MY14 haplotype was present in 21 of the 152 unique strains, with the remainder being of the N2 haplotype. Both haplotypes were found globally: the rarer MY14 haplotype was found in Europe, the United States, New Zealand and Chile. In almost all cases in which individuals with the MY14 haplotype were isolated, individuals of the N2 haplotype were isolated from proximal environments at the same time—a distribution that is consistent with balancing selection.

Sequence features of the *roam-1* region also fit the criteria for a region under balancing selection. The region encompassing *roam-1* had a relatively high Tajima's  $D$  of 1.01, a figure unusual both at the genomic level (at which  $<3.4\%$  of bins had a higher value) and in the centre of chromosome V where *srx-43* lies (at which  $<3.6\%$  of bins had a higher value). A phylogenetic analysis of the 152 strains revealed that the *roam-1* haplotype extends by approximately 30 kb before being disrupted by recombination events (Extended Data Fig. 7). Given the low outcrossing rate in wild-type *C. elegans* populations<sup>25,26</sup>, this short *roam-1* haplotype suggests co-occurrence of both haplotypes within interbreeding populations over many generations.

To assess directly the possibility that selection could act on *roam-1*, we designed competition experiments to compare the relative fitness of N2 and *roam-1<sub>MY14</sub>* strains. Experiments were conducted under high-density conditions to permit the accumulation and detection of endogenously produced *icas#9*, with competition applied by growing cultures past the point of starvation (that is, with limiting food). The first competition experiments were conducted on a simple lawn of *Escherichia coli* OP50 bacteria with a population of 20 N2 and 20



**Figure 5 | Bidirectional competitive selection at the *roam-1* locus.**

**a**, Diagram of 'boom-bust' competition experiments, with food depletion followed by 48 h of starvation before transfer. **b–d**, Competition on a 'simple lawn' showing allele ratio of DNA harvested at transfers 1 and 3. N2 versus *roam-1*<sub>MY14</sub> NIL (**b**), N2 *srx-43(lf)* versus *roam-1*<sub>MY14</sub> *srx-43(lf)* (**c**) and N2 *daf-22(lf)* versus *roam-1*<sub>MY14</sub> *daf-22(lf)*, without or with 10 nM exogenous icas#9 (**d**). **e**, 'Patchy lawn' competition between N2 and *roam-1*<sub>MY14</sub> NIL. Grey points represent individual competition experiments; red line indicates the mean. \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001 compared to an expected value of 0.5 by *t*-test with Bonferroni correction.

*roam-1*<sub>MY14</sub> age-matched adults (Fig. 5a). These conditions resulted in a growth advantage for the N2 genotype over *roam-1*<sub>MY14</sub> in the first cycle of competition that continued in subsequent cycles (Fig. 5b).

The two tested strains differ in the 182-kb *roam-1* region, which encompasses 81 genes. To investigate whether this competitive advantage required *srx-43*, the experiment was repeated using N2 *srx-43(lf)* and *roam-1*<sub>MY14</sub> *srx-43(lf)* strains. The competitive N2 advantage disappeared in this setting, indicating that the icas#9 receptor SRX-43 is essential for the competitive effect (Fig. 5c).

We assessed the role of endogenous pheromones by repeating the competition experiments with N2 and *roam-1*<sub>MY14</sub> strains with loss-of-function mutations in the gene *daf-22*, which is required for the secretion of ascarosides including icas#9. Mutations in *daf-22* eliminated the competitive advantage of the N2 strain over *roam-1*<sub>MY14</sub> (Fig. 5d). A partial advantage for N2 *daf-22(lf)* was recovered upon addition of exogenous 10 nM icas#9 (Fig. 5d). These results thus indicate that selection on the *roam-1* locus depends on pheromones.

The increased roaming of *roam-1*<sub>MY14</sub> animals at high population densities could be expected to cause greater exploration of a patchy food environment. In a second competition design, we seeded a patchy environment consisting of 16 small bacterial lawns with 20 N2 and 20 *roam-1*<sub>MY14</sub> adults (Fig. 5a). In these conditions, the N2 advantage was lost and instead a moderate but significant selection favoured *roam-1*<sub>MY14</sub> over N2 animals (Fig. 5e). Together, these results demonstrate that *roam-1* has a bidirectional effect on fitness that is dependent on *srx-43* expression, pheromone production and food distribution.

## Discussion

Conspecific individuals are informative elements of an animal's natural environment, in part because they represent competition for resources. Our results demonstrate that conspecific pheromones alter long-term foraging strategies and that natural variation in this behaviour stems from altered expression of the icas#9 receptor SRX-43. The complement of ascarosides produced by *C. elegans* varies with sex, age and feeding status<sup>27,28</sup>. The specificity of receptors such as SRX-43 provides

a mechanism through which this information can be detected by the nervous system, allowing it to regulate different behaviours and physiological responses. *srx-43* is expressed in ASI sensory neurons, which are targets of internal neuromodulators that regulate roaming and dwelling<sup>8</sup>, and represents a site of integration for internal and external influences on foraging behaviour. Although we do not know the suite of pheromones that are produced by *C. elegans* in the wild, the presence of secreted icas#9 in dense culture supernatants at concentrations 100-fold above those that suppress roaming suggest that it is a relevant regulator of foraging and that altered sensitivity to this molecule could affect animals' overall sensitivity to secreted ascarosides.

The *roam-1* QTL that encompasses *srx-43* has sequence features consistent with an area under balancing selection. While *srx-43* is an essential component of *roam-1*, it may not be the only gene in this QTL, or the only gene that is under balancing selection, as the haplotype extends for ~30 kb to include several other genes. Moreover, the behaviour identified here need not be the most important one in natural settings; it may represent one of several behavioural and physiological responses that facilitate adaptation to different environments. Balancing selection may be fairly common throughout the *C. elegans* genome: a recent report identified 61 highly divergent regions consistent with balancing selection that segregate among wild strains of *C. elegans*, including a second region 200 kb from *roam-1* (refs 24, 29; Extended Data Fig. 8). The composition of these regions is biased towards particular gene classes, including chemoreceptors, which may act as evolutionary hotspots.

At a conceptual level, behavioural genetics in animals, including humans, is dominated by evidence for gene–environment interactions<sup>30,31</sup>. Our results ground this abstraction in observation, showing that natural trait variation acts explicitly at the intersection of innate circuits and environment cues, with genetic changes allowing the differential incorporation of environmental information into innate foraging behaviours.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 27 March; accepted 14 September 2016.**

**Published online 31 October 2016.**

- Smith, J. M. The theory of games and the evolution of animal conflicts. *J. Theor. Biol.* **47**, 209–221 (1974).
- Dugatkin, L. A. & Reeve, H. K. *Game Theory and Animal Behavior*. (Oxford Univ. Press, 1998).
- Sokolowski, M. B., Pereira, H. S. & Hughes, K. Evolution of foraging behavior in *Drosophila* by density-dependent selection. *Proc. Natl Acad. Sci. USA* **94**, 7373–7377 (1997).
- Osborne, K. A. et al. Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. *Science* **277**, 834–836 (1997).
- Fitzpatrick, M. J., Feder, E., Rowe, L. & Sokolowski, M. B. Maintaining a behaviour polymorphism by frequency-dependent selection on a single gene. *Nature* **447**, 210–212 (2007).
- Fujiwara, M., Sengupta, P. & McIntire, S. L. Regulation of body size and behavioral state of *C. elegans* by sensory perception and the EGL-4 cGMP-dependent protein kinase. *Neuron* **36**, 1091–1102 (2002).
- Ben Arous, J., Laffont, S. & Chatenay, D. Molecular and sensory basis of a food related two-state behavior in *C. elegans*. *PLoS One* **4**, e7584 (2009).
- Flavell, S. W. et al. Serotonin and the neuropeptide PDF initiate and extend opposing behavioral states in *C. elegans*. *Cell* **154**, 1023–1035 (2013).
- Jeong, P.-Y. et al. Chemical structure and biological activity of the *Caenorhabditis elegans* dauer-inducing pheromone. *Nature* **433**, 541–545 (2005).
- Srinivasan, J. et al. A modular library of small molecule signals regulates social behaviors in *Caenorhabditis elegans*. *PLoS Biol.* **10**, e1001237 (2012).
- Macosko, E. Z. et al. A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature* **458**, 1171–1175 (2009).
- McGrath, P. T. et al. Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. *Neuron* **61**, 692–699 (2009).
- Greenberg, A. J., Moran, J. R., Coyne, J. A. & Wu, C.-I. Ecological adaptation during speciation revealed by precise gene replacement. *Science* **302**, 1754–1757 (2003).
- Andolfatto, P. et al. Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Res.* **21**, 610–617 (2011).



15. You, Y.-J., Kim, J., Raizen, D. M. & Avery, L. Insulin, cGMP, and TGF- $\beta$  signals regulate food intake and quiescence in *C. elegans*: a model for satiety. *Cell Metab.* **7**, 249–257 (2008).
16. Kim, K. *et al.* Two chemoreceptors mediate developmental effects of dauer pheromone in *C. elegans*. *Science* **326**, 994–998 (2009).
17. McGrath, P. T. *et al.* Parallel evolution of domesticated *Caenorhabditis* species targets pheromone receptor genes. *Nature* **477**, 321–325 (2011).
18. Schackwitz, W. S., Inoue, T. & Thomas, J. H. Chemosensory neurons function in parallel to mediate a pheromone response in *C. elegans*. *Neuron* **17**, 719–728 (1996).
19. Ren, P. *et al.* Control of *C. elegans* larval development by neuronal expression of a TGF- $\beta$  homolog. *Science* **274**, 1389–1391 (1996).
20. Akerboom, J. *et al.* Optimization of a GCaMP calcium indicator for neural activity imaging. *J. Neurosci.* **32**, 13819–13840 (2012).
21. Sorensen, P. W., Stacey, N. E. & Chamberlain, K. J. Differing behavioral and endocrinological effects of two female sex pheromones on male goldfish. *Horm. Behav.* **23**, 317–332 (1989).
22. Le Conte, Y. & Hefetz, A. Primer pheromones in social hymenoptera. *Annu. Rev. Entomol.* **53**, 523–542 (2008).
23. Thompson, O. *et al.* The million mutation project: a new approach to genetics in *Caenorhabditis elegans*. *Genome Res.* **23**, 1749–1762 (2013).
24. Thompson, O. A. *et al.* Remarkably divergent regions punctuate the genome assembly of the *Caenorhabditis elegans* Hawaiian strain CB4856. *Genetics* **200**, 975–989 (2015).
25. Barrière, A. & Félix, M.-A. High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Curr. Biol.* **15**, 1176–1184 (2005).
26. Barrière, A. & Félix, M. A. Natural variation and population genetics of *Caenorhabditis elegans*. *WormBook* <http://dx.doi.org/10.1895/wormbook.1.43.1> (2005).
27. Zhang, X. *et al.* Acyl-CoA oxidase complexes control the chemical message produced by *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **112**, 3955–3960 (2015).
28. Artyukhin, A. B. *et al.* Succinylated octopamine ascarosides and a new pathway of biogenic amine metabolism in *Caenorhabditis elegans*. *J. Biol. Chem.* **288**, 18778–18783 (2013).
29. Ghosh, R., Andersen, E. C., Shapiro, J. A., Gerke, J. P. & Kruglyak, L. Natural variation in a chloride channel subunit confers avermectin resistance in *C. elegans*. *Science* **335**, 574–578 (2012).
30. Bendesky, A. & Bargmann, C. I. Genetic contributions to behavioural diversity at the gene–environment interface. *Nat. Rev. Genet.* **12**, 809–820 (2011).
31. Kendler, K. S. *et al.* Stressful life events, genetic liability, and onset of an episode of major depression in women. *Am. J. Psychiatry* **152**, 833–842 (1995).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank E. Andersen for sharing the unpublished CeNDR database; the *Caenorhabditis* Genetics Center (NIH P40 OD010440) and the Million Mutation Project for strains; S. Flavell and A. Lopez for advice and insight; and P. Sengupta, M. O'Donnell, X. Jin and A. Sordillo for comments on the manuscript. R.A.B. and X.Z. were supported by the Research Corporation for Science Advancement (Cottrell Scholar Award, 22844). P.T.M. was supported by NIH grant R01GM114170 and the Ellison Medical Foundation. J.S.G. was supported by the NIH grant F30 MH101931-03. C.I.B. is an investigator of the HHMI. This work was supported by the Ellison Medical Foundation.

**Author Contributions** J.S.G. designed and performed the genetic, molecular biology, and behavioural experiments, together with M.B. for RIL analyses and I.G.I. for competition experiments. M.D. performed calcium imaging experiments. E.Z.M. discovered the effect of pheromones on foraging. X.Z. and R.A.B. analysed pheromone production and synthesized pure pheromones. D.J.C. and P.T.M. performed population genetic analysis. J.S.G., P.T.M. and C.I.B. analysed and interpreted data. J.S.G. and C.I.B. wrote the manuscript, with input from all authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.I.B. ([cori@rockefeller.edu](mailto:cori@rockefeller.edu)).

**Reviewer Information** *Nature* thanks E. Haag, S. Lockery and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

**Nematode culture.** All strains were grown at 21–22 °C on nematode growth-medium plates seeded with *E. coli* OP50 bacteria<sup>32</sup>. For OP50 cultures, a single colony was inoculated into 100 ml of LB and grown for 48 h at 21–22 °C. Transgenic lines were generated by standard injection methods and included the desired transgene, a fluorescent co-injection marker and an empty vector, bringing the total DNA concentration up to 100 ng  $\mu\text{l}^{-1}$ . For each transgene, three independent extrachromosomal lines that propagated the transgene at high rates were tested in parallel to account for variability typical of such strains. All mutagenized strains were back-crossed 5–7 times before characterization.

**Natural isolates and origin of wild-type strains.** N2 (Bristol, UK), CX12311 *kyIR1* (chrV, CB4856 > N2) V; *qqIR1* (chrX, CB4856 > N2), CB4856 (Hawaii, USA), JU258 (Riberio Frio, Madeira), MY1 (Lingen, Germany), MY14 (Mecklenbeck, Germany), JU775 (Lisbon, Portugal), JU1400 (Sevilla, Spain), JU1652 (Montevideo, Uruguay), AB1 (Adelaide, Australia), MY16 (Mecklenbeck, Germany), JU1171 (Concepcion, Chile), MY6 (Roxel, Germany), JU360 (Franconville, France), ED3021 (Edinburgh, Scotland), MY2 (Roxel, Germany).

Strain CX12311 bears ancestral alleles of the *npr-1* and *glb-5* genes, which affect oxygen sensitivity and are mutated in the N2 laboratory strain<sup>17</sup>; it is therefore used as a comparison strain for wild-type strains bearing the ancestral alleles.

**MY14–CX12311 RILs.** CX14697–CX14712, CX14731–CX14748, CX14750–CX14757, CX14783, CX14784, CX14786–CX14820, CX14822–CX14839. Genotypes inferred from low-coverage genomic sequence and behavioural data are included as Supplementary Table 1.

**Near-isogenic lines.** CX15881 *kyIR142* (chrV:~14.3–16.8 Mb, CX12311 > MY14), CX15878 *kyIR139* (chrV:~14.3–16.8 Mb, MY14 > CX12311), CX15883 *kyIR144* (chrV:~14.3–16.8 Mb, MY14 > N2), CX16075 *kyIR147* (chrV:~15.861–16.8 Mb, MY14 > N2), CX16140 *kyIR153* (chrV:~16.043–16.8 Mb, MY14 > N2), CX16300 *kyIR163* (chrV:~15.861–16.043 Mb, MY14 > N2; *roam-1*<sub>MY14</sub>), CX16294 *kyIR157* (chrV:~15.861–16.006 Mb, MY14 > N2).

**Transgenic lines.** DNAs are N2-derived unless otherwise noted. CX16884 *kyIR163* V; *kyEx5851* (*Psrx-43*<sup>N2</sup>::*srx-43*<sup>N2</sup>::*sl2::GFP*, 2.5 ng  $\mu\text{l}^{-1}$ ; *Pmyo3::mcherry*, 5 ng  $\mu\text{l}^{-1}$ ), CX17202 *kyIR163* V; *kyEx6012* (*Psrx-43*<sup>N2</sup>::*srx-43*<sup>N2</sup>(*nonsense*)::*sl2::GFP*, 2.5 ng  $\mu\text{l}^{-1}$ ; *Pmyo3::mcherry*, 5 ng  $\mu\text{l}^{-1}$ ), CX16881 *srx-43*(*gk922634*) V; *kyEx5848* (*srx-43*<sup>N2</sup>, 2.5 ng  $\mu\text{l}^{-1}$ ; *Pmyo3::mcherry*, 5 ng  $\mu\text{l}^{-1}$ ); *gk922634* changes R160 to an opal stop codon, CX17204 *kyEx6013* (*Psrx-43*<sup>N2</sup>::*srx-43*<sup>N2</sup>::*GFP*, 50 ng  $\mu\text{l}^{-1}$ ; *Pelt-2::GFP*, 5 ng  $\mu\text{l}^{-1}$ ), CX16943 *kyIR163* V; *kyEx5894* (*Psrx-43*<sup>MY14</sup>::*srx-43*<sup>MY14</sup>::*sl2::GFP*, 2.5 ng  $\mu\text{l}^{-1}$ ; *Pmyo3::mcherry*, 5 ng  $\mu\text{l}^{-1}$ ), CX16425 *kyIs602* (*Psra-6::GCAMP3.0*, 75 ng  $\mu\text{l}^{-1}$ ; *Pcoel::GFP*, 10 ng  $\mu\text{l}^{-1}$ ); *kyEx5594* (*Psra-6::srx-43*<sup>N2</sup>, 50 ng  $\mu\text{l}^{-1}$ ; *Pmyo3::mcherry*, 5 ng  $\mu\text{l}^{-1}$ ), CX16931 *kyIs602*; *kyEx5885* (*Psra-6::srx-43*<sup>MY14</sup>, 50 ng  $\mu\text{l}^{-1}$ ; *Pmyo3::mcherry*, 5 ng  $\mu\text{l}^{-1}$ ), CX17196 *kySi66* (*MosSCI Psrx-43*<sup>N2</sup>::*srx-43*<sup>N2</sup>) II; *srx-43*(*gk922634*) V, outcrossed 4 $\times$ , CX17198 *kySi68* (*MosSCI Psrx-43*<sup>MY14</sup>::*srx-43*<sup>MY14</sup>) II; *srx-43*(*gk922634*) V, outcrossed 4 $\times$ , CX17201 *kySi71* (*MosSCI Psrx-43*<sup>N2</sup>::*srx-43*<sup>MY14</sup>) II; *srx-43*(*gk922634*) V, outcrossed 4 $\times$ , CX17203 *kySi72* (*MosSCI Psrx-43*<sup>MY14</sup>::*srx-43*<sup>N2</sup>) II; *srx-43*(*gk922634*) V, outcrossed 4 $\times$ , FK181 *ksIs2* (*Pdaf-7::GFP + rol-6*(*su1006*)), CX16958 *kyIR163* V; *ksIs2*

**Mutants.** CX16849 *srx-43*(*gk922634*) V, outcrossed 5 $\times$  to N2. *gk922634* is R160opal. This mutation was provided by the Million Mutation Project<sup>23</sup>.

CX16935 *kyIR163 srx-43*(*ky1019*) V. *ky1019* is a CRISPR/Cas9-induced indel mutation that causes a frame-shift mutation after the first transmembrane domain (insertion (TCACTGAGTTCGAAT), deletion (CCCCG), final sequence TCGCAGCTCTCAAGTTCAGTTCGAATTCGGAATTCTC; insertion is underlined). We used the coCRISPR protocol described previously<sup>33</sup>. Young adults were injected with a mix of plasmids containing Cas9, guide RNA targeting *rol-6*, and guide RNA targeting the location of the desired mutation, as well as a ssDNA template for inducing a dominant *rol-6*(*su1006*) mutation. F1 animals with a roller phenotype were isolated and allowed to lay eggs before secondary screening for the target mutation by Sanger sequencing.

Previously described mutants included JT5464 *daf-7*(*e1372*) III; *daf-3*(*e1376*) X, CX17307 *daf-7*(*e1372*) III; *kyIR163* V; *daf-3*(*e1376*) X, CX13846 *daf-22*(*ok693*) II, CX17082 *daf-22*(*ok693*) II; *kyIR163* V.

**Behavioural analysis.** Exploration assays<sup>8</sup> were conducted on 35-mm Petri dishes evenly seeded with 100  $\mu\text{l}$  of OP50 *E. coli* bacteria 24 h before the start of the assay. Individual two-day-old L4 hermaphrodites were placed in the centre of the plate. After 16 h, plates were placed on a grid containing 86 squares and the number of full or partial squares that contained tracks was quantified by an investigator blinded to the genotype. Pheromones or control solvent were mixed into the agar. A pheromone response for each animal on an ascaroside plate was determined with respect to the behaviour of control animals that were tested on ascaroside-free plates on the same day.

Individual pheromone response was defined as the mean number of squares entered by controls tested on the same day, subtracted by the number of squares

entered by the test animal. Group pheromone response was defined as the mean pheromone response of all individuals tested across days. For statistical analysis, *n* was defined as the total number of animals in the ascaroside group.

N2-derived strains were tested in 21% oxygen (Figs 1, 2e, f, 3, 5 and Extended Data Figs 1, 5). All naturally isolated strains and CX12311-derived strains bearing ancestral alleles of *npr-1* and *glb-5* were tested in 8% oxygen to suppress the oxygen-dependent roaming behaviour of ancestral *npr-1* alleles<sup>12,34</sup> (Figs 2a, b, d, 4 and Extended Data Fig. 2).

Direct examination of roaming and dwelling was modified from previously reported techniques<sup>8</sup>. At 14.5 h before testing, 25 L4 larvae were picked and placed onto 150-mm test plates thinly seeded with 1.5 ml of OP50 bacteria with or without synthetic pheromone. Video recording was conducted under red light to minimize behavioural response to imaging conditions. 1.5-h-long videos were recorded at 3 frames  $\text{s}^{-1}$  using Streampix software (Norpix Inc.) and a 6.6 MP PL-B781F CMOS camera (PixeLINK). Custom Matlab scripts<sup>8</sup> were employed to determine worm trajectories and conduct a two-state hidden Markov model determining the most probable state path for each animal and thereby measure roaming- and dwelling-state durations.

The low basal exploration rate in *daf-7*(*lf*) mutants<sup>7</sup> prevented a direct assessment of the effect of *icas#9* on foraging behaviour. Instead, we examined *daf-7*(*lf*) *daf-3*(*lf*) double mutants, as *daf-7* canonically acts by antagonizing *daf-3*, which encodes a co-SMAD transcriptional regulator. We found that *daf-3* mutations suppressed the low basal exploration rate of *daf-7* mutants. N2 *daf-7*(*lf*) *daf-3*(*lf*) animals explored control plates to a greater degree than the wild type, so larger (10 cm) exploration assay plates were used to score these strains.

**Statistics.** No statistical methods were used to predetermine sample size.

Most experiments were repeated on three separate days. For exploration assays, the standard group size on a single day was six; this ensured sufficient power to detect moderate effects, while also limiting the influence of daily variation. All plates with a healthy adult animal at the end of the assay were scored and included in the analysis. Randomization was ensured using the following or similar approach: at the start of each exploration assay, six animals were placed on a pick at a time. They were then transferred individually to three control plates and then to three *icas#9* plates in the order the animals came off the pick. Assays were scored by an experimenter blinded to the condition or genotype.

Most statistical comparisons were performed using ANOVA with Dunnett's correction for multiple comparisons or using a (two-sided) *t*-test, as noted in the figure legends. The normality of the data was tested with D'Agostino–Pearson omnibus test. Bartlett's test was used to check for differences in variance between groups being statistically compared. N2 groups in Fig. 3c and Extended Data Figure 5e did not pass a normality test. As the *n* number was large (>15), ANOVA was still an appropriate test. Moreover, the findings were still significant when a nonparametric test was used (Kruskal–Wallis with Dunn multiple comparison test).

**Ascaroside quantification.** We grew 150 ml unsynchronized worm cultures for 9 days and fed on *E. coli* (HB101 or OP50), as described<sup>27</sup>. Extracts were generated from the culture medium and analysed by liquid chromatography–tandem mass spectroscopy (LC–MS/MS), as described previously<sup>27</sup>, and analysed on a Thermo Scientific TSQ Quantum Access MAX, with the collision gas pressure set to 1 mTorr. Ascaroside concentrations present in the culture were quantified using the corresponding synthetic standards, with the following exceptions: synthetic *ascr#18* was used to quantify *ascr#22* and *ascr#26* and synthetic *icas#3* was used to quantify *icas#1* and *icas#10*.

**Recombinant inbred lines.** The MY14–CX12311 recombinant inbred lines were generated by crossing MY14 males with CX12311 hermaphrodites and CX12311 males with MY14 hermaphrodites, to ensure the mitochondrial DNA from both strains were equally represented in the RILs. In total, 94 F2 animals were individually picked, placed onto plates and inbred through self-fertilization for 10 generations. RIL genotyping was conducted by low-coverage shotgun sequencing<sup>14</sup>. Genomic DNA was fragmented and attached to sequencing adapters with a Nextera DNA Library Prep Kit (Illumina). Samples were pooled and sequenced on an Illumina HiSeq 2000. Sequencing reads from each strain were mapped to the WS235 release of the *C. elegans* genome using the Burrows–Wheeler Aligner to create bam files for further analysis<sup>35</sup>. The set of MY14/N2 single nucleotide variants (SNVs) identified in the Million Mutation Project were used for genotyping purposes<sup>23</sup>. Each genetic variant was genotyped in each strain. Owing to the low coverage, the majority of SNVs were not genotyped. To improve the data coverage, we grouped 200 neighbouring SNV genotypes together to create a consensus genotype for 540 bins (either N2, MY14 or heterozygous). These genotypes were used for QTL mapping.

**QTL mapping.** The pheromone response index was used as the phenotype in combination with the 540 genotype bins from above. R/qtl was used to perform a one-dimensional scan using marker regression on all 540 markers. The significance

threshold was determined using 1,000 permutation tests. The effect-size of the *roam-1* locus was estimated using the fitqtl function with a single QTL. The peak of the *roam-1* locus (chromosome V: 16,451,686–16,579,457) was used as an additive and interactive covariate for additional one-dimensional scans, assuming a normal model. The significance threshold for these two tests was also determined using 1,000 permutation tests.

**NIL mapping.** Before the detailed QTL mapping by sequencing described above, the *roam-1* QTL was localized to 2.5 Mb (chrV: 14.3–16.8 Mb) by examining 14 high-confidence phenotypically extreme RILs (Supplementary Table 1). This result, which was confirmed by the full analysis, guided the initial generation of NILs. The NIL *kyIR142* was produced by backcrossing the RIL CX14816 nine times to MY14, maintaining N2 alleles at chrV: 14.3 and chrV: 16.8 Mb at each generation. The NIL *kyIR139* was produced by backcrossing the RIL CX14708 nine times to CX12311, maintaining MY14 alleles at chrV: 14.3 and chrV: 16.8 Mb. The NIL: *kyIR144* was produced by crossing *kyIR139* with N2 and isolating recombinants with the N2 allele of *glb-5* (chrV: 5.56 Mb), the MY14 alleles at chrV: 14.3 and 16.8 Mb, and the N2 allele of *npr-1* on chrX. The NILs *kyIR147* and *kyIR153* were created by crossing *kyIR144* with N2 and identifying progeny with the N2 allele at chrV: 14.3 Mb and the MY14 allele at chrV: 16.8 Mb.

**High-density recombination mapping.** We crossed *kyIR147* with males from CX16290, a N2 strain with an integrated fluorescent marker at chrV: 15.83 Mb. F1 progeny were identified by fluorescence, picked to growth plates, and allowed to lay eggs for 12 h. Following 3 days of growth, ~2,600 non-fluorescent F2 animals were sorted individually into wells of 96-well plates by a worm sorter (COPAS Biosort Systems, Union Biometrika). These F2 were grown in 200 µl of S Basal buffer (5.85 g NaCl, 1 g K<sub>2</sub> HPO<sub>4</sub>, 6 g KH<sub>2</sub>PO<sub>4</sub>, 5 mg cholesterol per litre) with cholesterol, supplemented with OP50 bacteria. A fraction of the F3 progeny from each isolate were lysed and genotyped at chrV: 16.069 Mb. Those with an N2 allele at chrV: 16.069 Mb were genotyped at chrV: 15.861 Mb. Twelve recombinants with an N2 allele at chrV: 16.069 Mb and a MY14 allele at chrV: 15.861 Mb were isolated and characterized behaviourally, among which were *kyIR163* and *kyIR157* (Fig. 2f). The N2 NIL with *kyIR163* (182 kb of MY14 sequence) is referred to as *roam-1<sub>MY14</sub>*.

**Imaging.** Calcium imaging experiments were performed and analysed as described previously<sup>36</sup>. In brief, young adult animals were placed into custom-made 3 mm<sup>2</sup> microfluidic polydimethylsiloxane devices that permit rapid changes in stimulus solution. Each device contains two arenas, allowing for simultaneous imaging of two genotypes with approximately ten animals each. Animals were transferred to the arenas in S-Basal buffer and paralyzed for 80–100 min in 1 mM (–)–tetramisole hydrochloride. Experiments consisted of four pulses of 10 s of stimulus separated by 30 s of buffer, with an additional 60 s between stimulus types. Tiff stacks were acquired at 10 frames s<sup>–1</sup> at 5× magnification (Hamamatsu Orca Flash 4 sCMOS), with 10 ms pulsed illumination every 100 ms (Sola, Lumencor; 470/40 nm excitation).

Fluorescence levels were analysed using a custom ImageJ script that integrates and background-subtracts fluorescence levels of the ASH neuron cell body (4 × 4 pixel region of interest). Using MATLAB, the calcium responses were normalized for each stimulus type by dividing fluorescence levels by the baseline fluorescence, defined as the average fluorescence of the 10 s preceding the first pulse of the stimulus. Each experiment was performed a total of four times over two separate days. Animals were pooled together by strain to calculate population mean and standard error (N2 *srx-43* allele, 23 animals; MY14 *srx-43* allele, 30 animals; array negative control, 19 animals). Experiments were conducted on two days.

For GFP expression studies, live adult animals were mounted on 2% agarose pads containing 5 mM sodium azide. Images were collected with a 100× objective on a Zeiss Axio Imager.Z1 Apotome microscope with a Zeiss AxioCam MRm CCD camera. For *daf-7* reporter studies, expression was quantified 16–24 h after L4 animals were placed on exploration assay plates. Images were processed in Metamorph and ImageJ to generate a maximum-intensity Z-projection. Reporter values were assessed as the mean grey value for a 16-pixel-radius circle centred over the cell body minus the mean background intensity. Both ASI neurons were analysed in each animal; experiments were performed over three days.

**Digital PCR.** Digital PCR was conducted on a QuantStudio 3D digital PCR platform (Thermo Fisher), and analysed on the QuantStudio 3D AnalysisSuite Cloud.

The *srx-43* mRNA expression studies were conducted on synchronized L4 worms 48 h after laying. RNA was collected on RNeasy Mini columns (Qiagen) and treated with DNase (Qiagen). SuperScript III First-Strand Synthesis System (Thermo Fisher) was used to create cDNA libraries. Custom TaqMan Expression Assays (Thermo Fisher) were used for *srx-43* quantification, and the tubulin gene, *tbb-1*, was used for normalization of digital PCR.

For quantitative analysis of the competition experiments, DNA was extracted with a standard phenol–chloroform protocol. Custom TaqMan SNP Genotyping Assays (Thermo Fisher) were used to determine the relative ratio of N2 versus *roam-1<sub>MY14</sub>* DNA by digital PCR. The assay was validated with known ratios of N2 to *roam-1<sub>MY14</sub>* DNA (Extended Data Fig. 9).

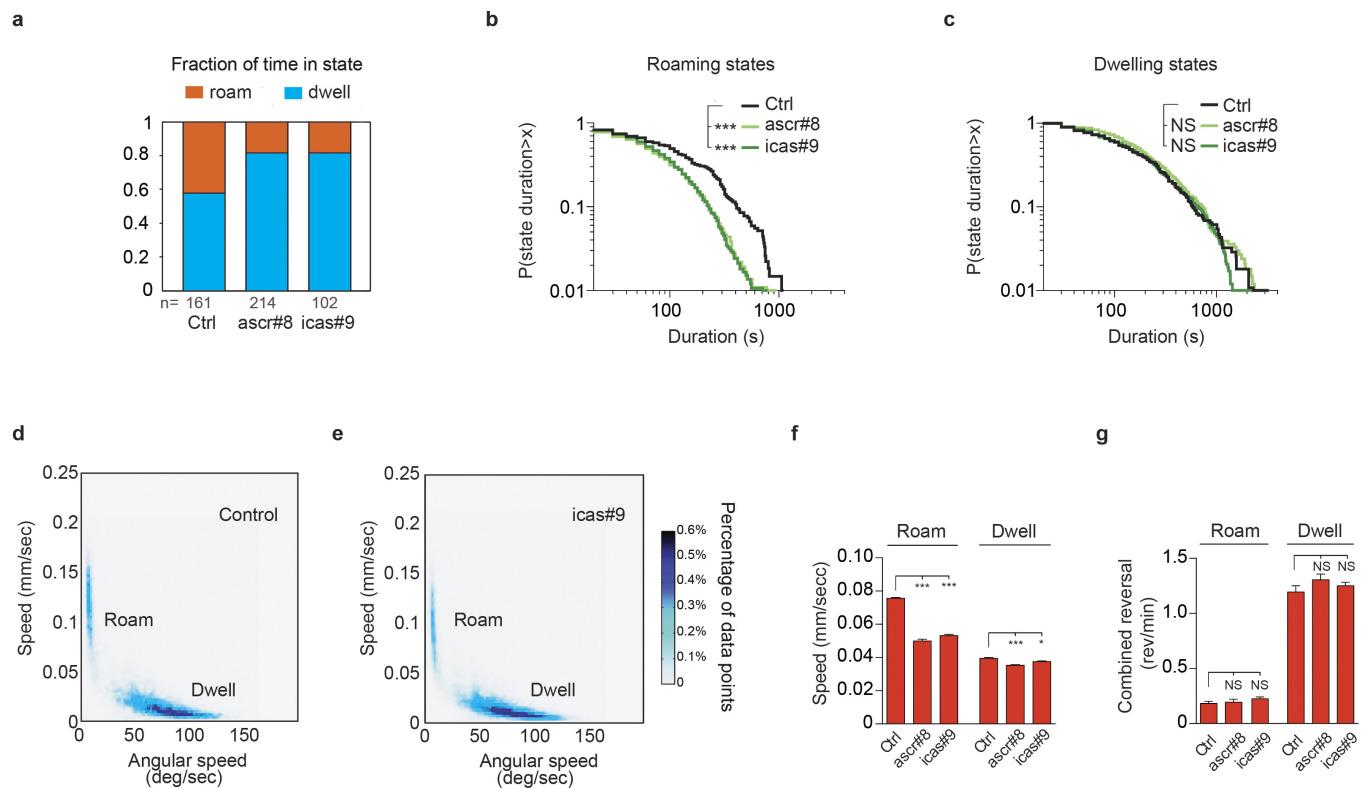
**Population genetics.** To create the gene and organism phylogenies, we used SNV data downloaded from the Million Mutation Project (<http://genome.sfu.ca/mmp/>) or the CeNDR resource (<http://www.elegansvariation.org>). For the CeNDR dataset, MY14 was assumed to be clonal or near-clonal with MY23, as was suggested by RAD sequencing. Software was written in Python using the Biopython module to create a neighbour joining tree. For the *roam-1* locus, SNVs on chrV between 16,010,000 and 16,030,000 were used. For the *glc-1* locus, SNVs on chrV between 16,181,000 and 16,222,000 were used. All SNVs were used to construct the whole-genome strain tree. Number of genetic variants and Tajima's *D* were calculated on 5-kb bins using vcftools<sup>37</sup>. *d<sub>N</sub>/d<sub>S</sub>* was calculated by counting using custom Python scripts analysing variants between MY23 and the N2 reference. Phylogenies of *srx-43* and closely related genes were performed using protein sequences obtained previously<sup>38</sup>.

**Fitness assays.** Competition experiments consisted of three boom–bust cycles. During the boom phase, population growth led to the rapid depletion of food, initiating the bust phase, which lasted for two days. Simple lawn competition experiments were conducted on 100-mm NGM agar plates with a single lawn formed from 800 µl of saturated OP50 culture. Patchy lawn competition experiments were conducted on 150-mm NGM agar plates with a 200 µl ring-shaped OP50 lawn in the centre of the plate surrounded by 15 small 40-µl lawns (Fig. 5a); at the assay start and at transfers animals were placed in the centre of the plate.

Populations were initiated from 20 N2-type and 20 *roam-1<sub>MY14</sub>*-type age-synchronized young adult animals. The initial population depleted food within 4 days, and on day 6 animals were washed into M9 media. 20% of the suspension was transferred to a new plate and the remainder was lysed for quantitative DNA analysis. For the second and third boom–bust cycle, food resources were depleted in 2 days and the plates were kept starved for an additional 2 days. Following the second bust phase, 20% of the animals were transferred to a new plate; following the third bust phase, the entire population was harvested for DNA extraction.

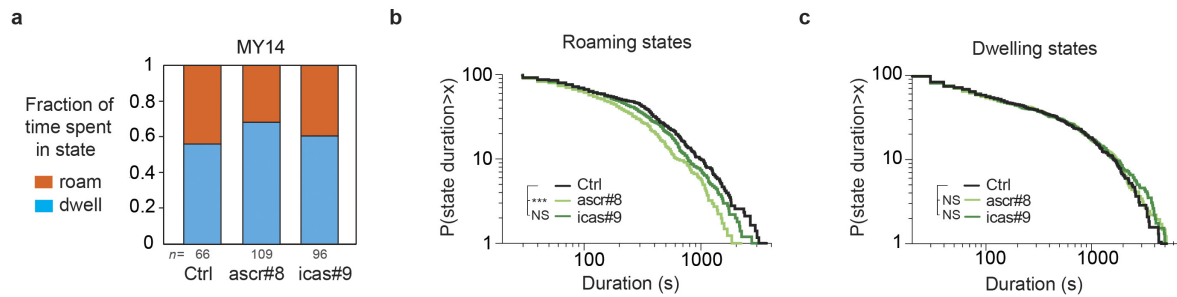
32. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
33. Arribere, J. A. et al. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics* **198**, 837–846 (2014).
34. Cheung, B. H. H., Cohen, M., Rogers, C., Albayram, O. & de Bono, M. Experience-dependent modulation of *C. elegans* behavior by ambient oxygen. *Curr. Biol.* **15**, 905–917 (2005).
35. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
36. Larsch, J., Ventimiglia, D., Bargmann, C. I. & Albrecht, D. R. High-throughput imaging of neuronal activity in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **110**, E4266–E4273 (2013).
37. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
38. Thomas, J. H. & Robertson, H. M. The *Caenorhabditis* chemoreceptor gene families. *BMC Biol.* **6**, 42 (2008).





**Extended Data Figure 1 | Roaming and dwelling states in the presence of ascariosides.** **a**, Roaming and dwelling behaviours scored from video analysis.  $n = 102$ – $214$  tracks per data point. **b**, **c**, Cumulative distribution of roaming (**b**) and dwelling (**c**) state durations for animals in **a**. \*\*\* $P < 0.001$  by log-rank test; ns, not significant. **d**, **e**, Scatter plot of average speed and angular speed (a measure of turning rate) in 10 s intervals taken from 1.5-h-long video recordings of wild-type animals

in control (**d**) and icas#9 (**e**) conditions. Roaming animals move quickly and turn infrequently compared with dwelling animals. Note the bimodal distribution defining distinct behavioural states. Control, 161 tracks; icas#9, 102 tracks. **f**, **g**, Speed following a reversal (**f**) and the reversal rate (**g**) for roaming or dwelling animals. Roaming speed is slightly slower in ascariosides (**e**, **f**). Data presented as mean  $\pm$  s.e.m. \* $P < 0.05$ , \*\*\* $P < 0.001$  by ANOVA with Dunnett correction; ns, not significant.

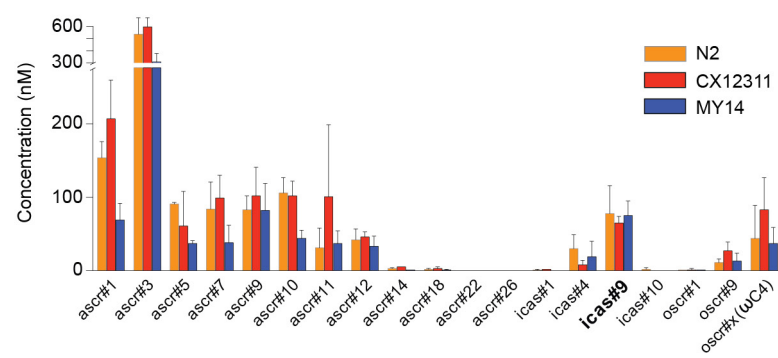


### Extended Data Figure 2 | Roaming and dwelling behaviour of MY14.

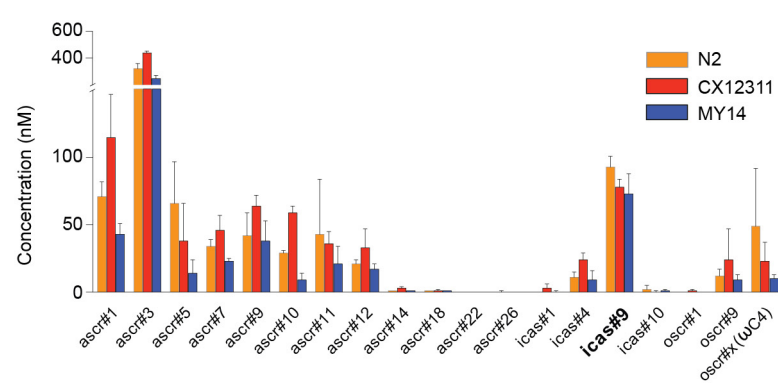
**a**, Fraction of time that MY14 animals spend roaming or dwelling in control, ascr#8, and icas#9 conditions;  $n = 66$ –109 tracks per data point. Assays were conducted in 8% O<sub>2</sub>. **b**, **c**, Cumulative distribution of roaming (**b**) and dwelling (**c**) state durations for MY14 animals scored in **a**.

Roaming states are significantly shorter in the presence of ascr#8 ( $t_{1/2} = \sim 150$  s, versus  $\sim 220$  s in controls), but are not significantly affected by icas#9 ( $t_{1/2} = \sim 190$  s). Roaming states may also be longer at baseline in MY14 than in N2 (see Extended Data Fig. 1). \*\*\* $P < 0.001$  by log-rank test.

a

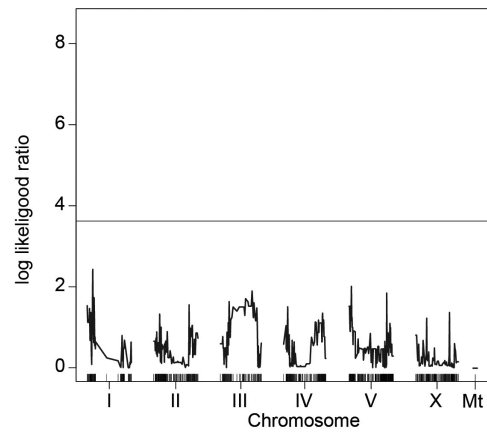
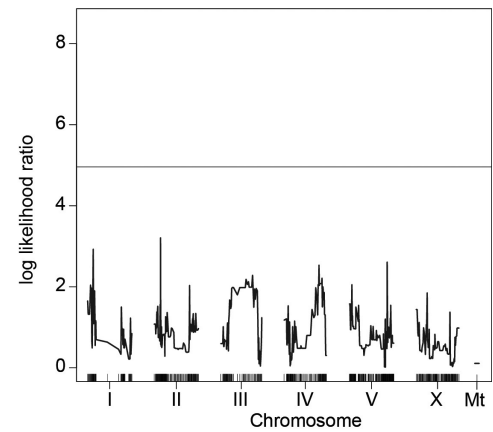


b

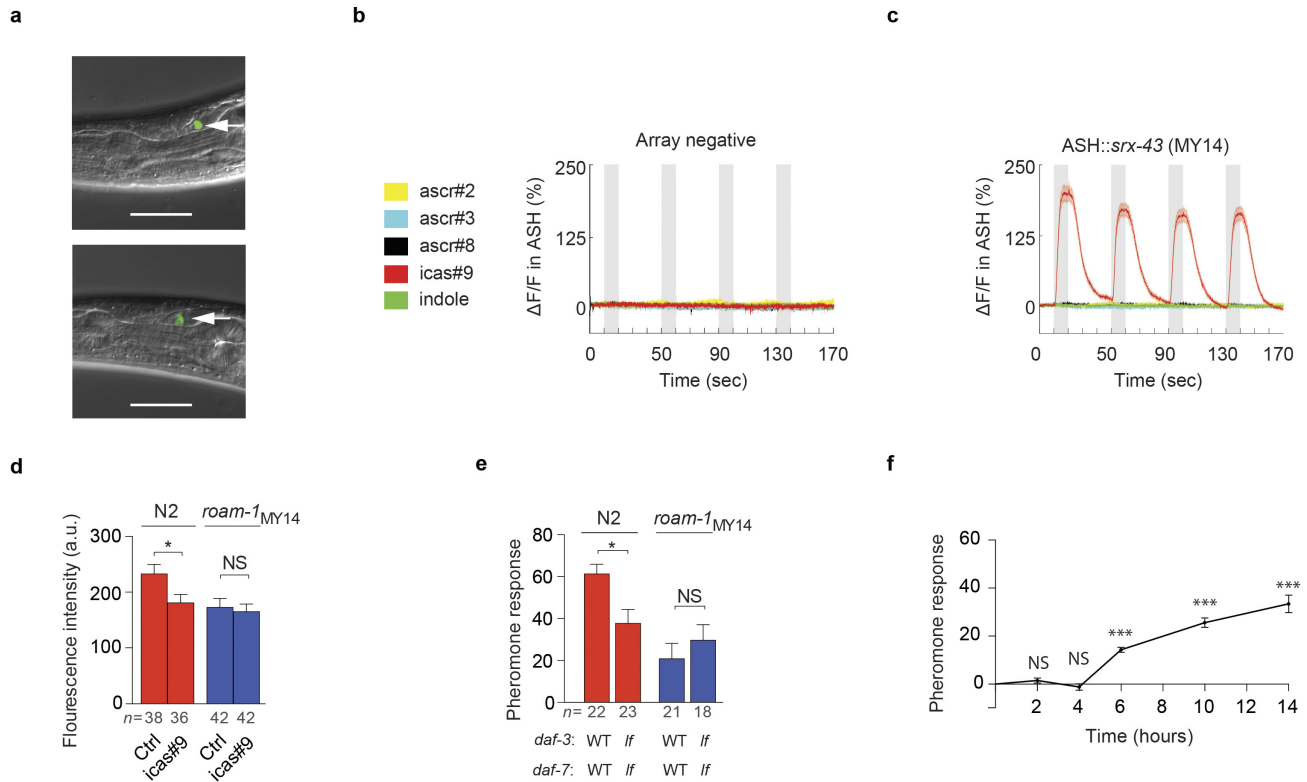


**Extended Data Figure 3 | Ascarosides produced by wild-type strains.** **a, b,** LC–MS/MS analysis of ascarosides secreted by N2, CX12311 and MY14 strains grown on OP50 (**a**) or HB101 (**b**) bacteria. icas#9 is produced at similar levels by icas#9-sensitive and icas#9-resistant strains.  $n = 2$  (**a**) or 3 (**b**) culture extracts per genotype.



**a****b**

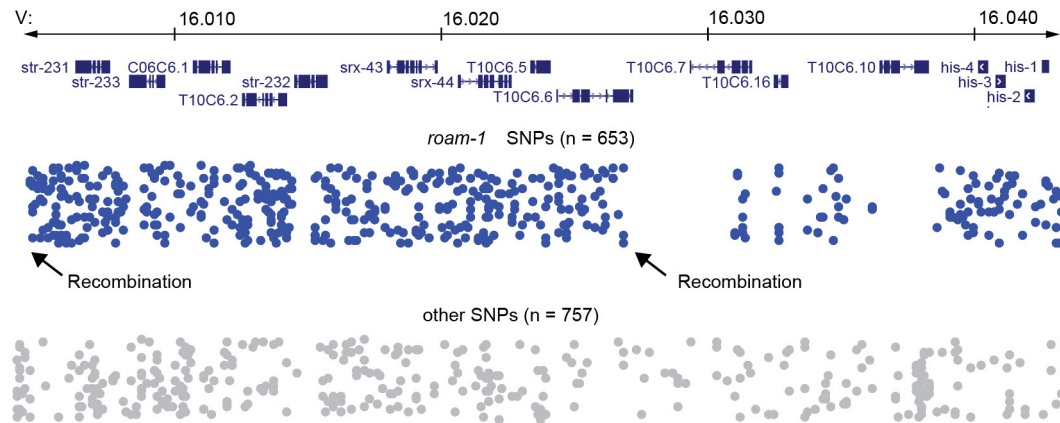
**Extended Data Figure 4 | Covariate analysis of 94 RILs. a, b,** Covariate analysis controlling for the *roam-1* genotype, testing for additive (a) or interactive (b) QTLs at other loci. The horizontal line denotes the  $P < 0.05$  genome-wide significance threshold. LOD, log likelihood ratio.



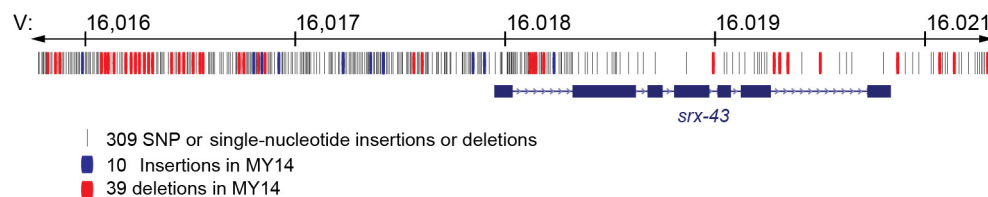
**Extended Data Figure 5 | Signal transduction by SRX-43.** **a**, Expression of *Psrx-43::srx-43::SL2::GFP* bicistronic reporter transgenes bearing N2 (top) or MY14 (bottom) *srx-43* sequences. Arrows indicate cell bodies of ASI sensory neurons. Scale bars, 50  $\mu$ m. **b**, ASH sensory neurons are insensitive to multiple ascarosides. ASH calcium imaging with GCaMP3 in control animals that do not express the *srx-43* transgene, isolated as non-transgenic siblings of transgenic animals tested in Fig. 3f ( $n = 19$ ). Ascarosides tested at 10 nM. **c**, SRX-43 from MY14 confers icas#9 sensitivity on ASH neurons. Compare SRX-43 from N2 in Fig. 3f. **d**, icas#9 decreases *daf-7::GFP* expression in ASI neurons of N2 but not

*roam-1<sub>MY14</sub>* adults. Bars indicate mean fluorescence intensity  $\pm$  s.e.m.  $*P < 0.05$  by ANOVA with Tukey's multiple comparisons test. **e**, Responses to icas#9 of *daf-7(lf)* mutants are attenuated in N2 but not in *roam-1<sub>MY14</sub>* genetic backgrounds. Modified exploration assays were conducted on strains including *daf-3(lf)* alleles (see methods).  $*P < 0.05$  by *t*-test. Data presented as mean  $\pm$  s.e.m. **f**, Time course for icas#9 response in exploration assay. Pheromone response expressed as mean  $\pm$  s.e.m. for 2, 4, 6, 10, and 14 h following initiation of exploration assay.  $***P < 0.001$  by *t*-test with Bonferroni correction comparing squares entered in control plates versus 10 nm icas#9 plates;  $n = 12$  for all time points.

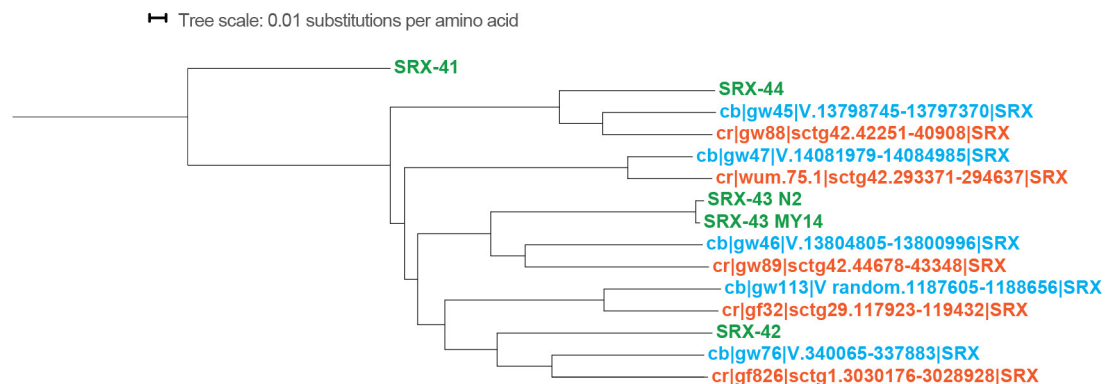
a



b



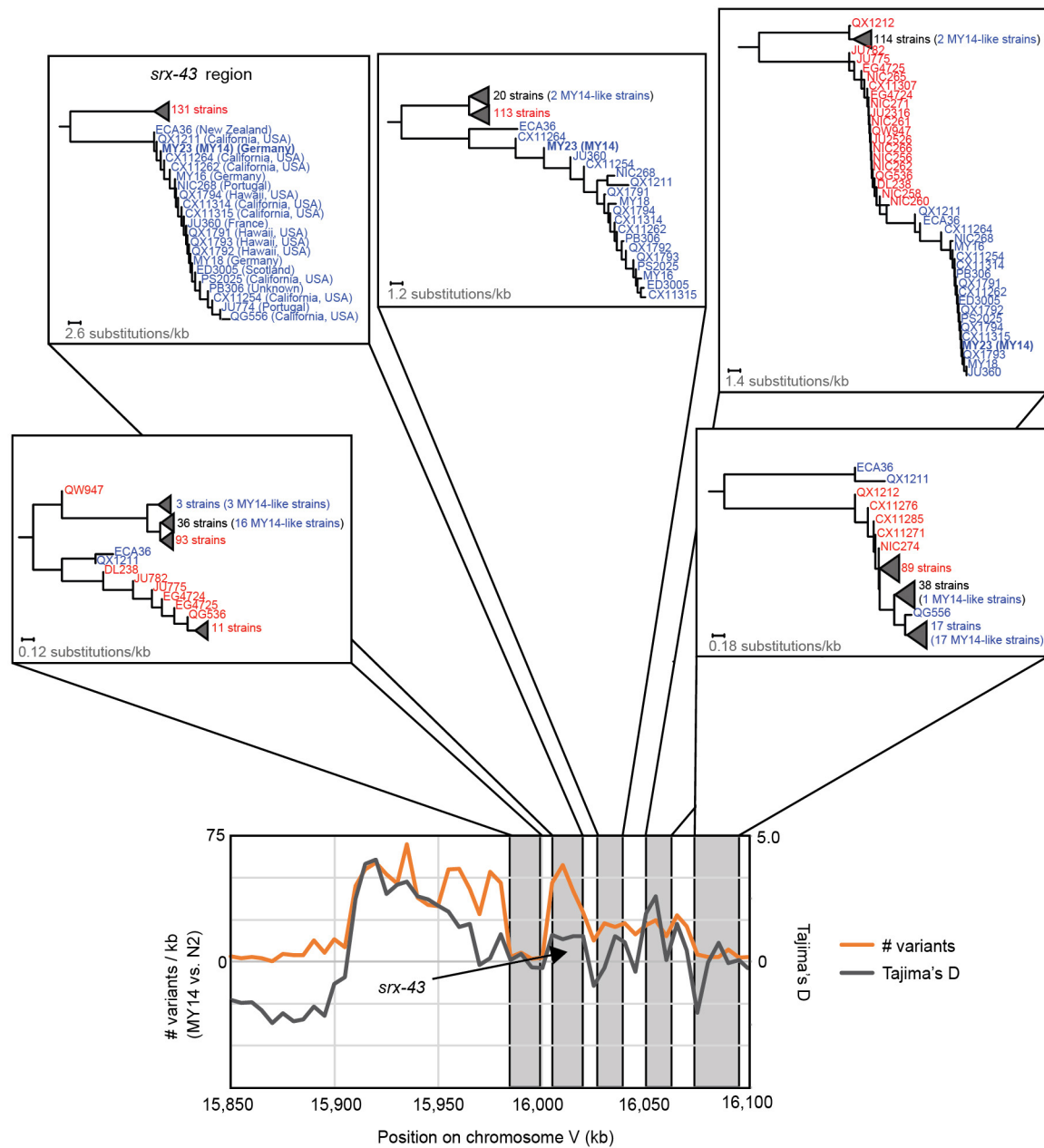
c



**Extended Data Figure 6 | Alternative *roam-1* alleles have high sequence variability.** **a**, The *roam-1* QTL region (top). *roam-1* SNPs are SNPs, when compared to the N2 reference genome, that are shared by JU360, MY2, MY14, ED3021, JU1171, MY16 and MY6 and not by any other strains, according to the Million Mutation Project. This defines the *roam-1*<sub>MY14</sub> haplotype. Other SNPs denote all other SNPs with respect to the N2 reference genome found in any of the 40 wild isolates in the Million Mutation Project. **b**, Polymorphisms of the *srx-43* promoter and coding region revealed by Sanger sequencing. Despite the high rate of polymorphism, there are only four non-synonymous mutations in the MY14 coding sequence detected by Sanger sequencing; three of these four were detected by the Million Mutation Project (Extended Data Table 1). We confirmed that the MY14 and N2 sequences are alleles of the same

gene by examining sequence reads of the MY14-like strain MY23 in the CeNDR data set (<http://www.elegansvariation.org>) and aligning each read to N2 and MY14 sequence for the *srx-43* region as determined by Sanger sequencing. We observed that 7,272 of the MY23 (MY14) reads better matched the MY14 Sanger sequence and 4 of the reads better matched the N2 reference sequence, as would be expected if MY14 and N2 each bear one alternative allele of the gene. **c**, Phylogeny constructed for *srx-43* and related genes in *C. elegans*, *C. briggsae* and *C. remanei* demonstrates that the *srx-43* alleles in N2 and MY14 are closely related alleles of a single gene. Genes are colour-coded by species (green, *C. elegans*; blue, *C. briggsae*; orange, *C. remanei*). Protein sequences and gene names are as previously described<sup>38</sup>.

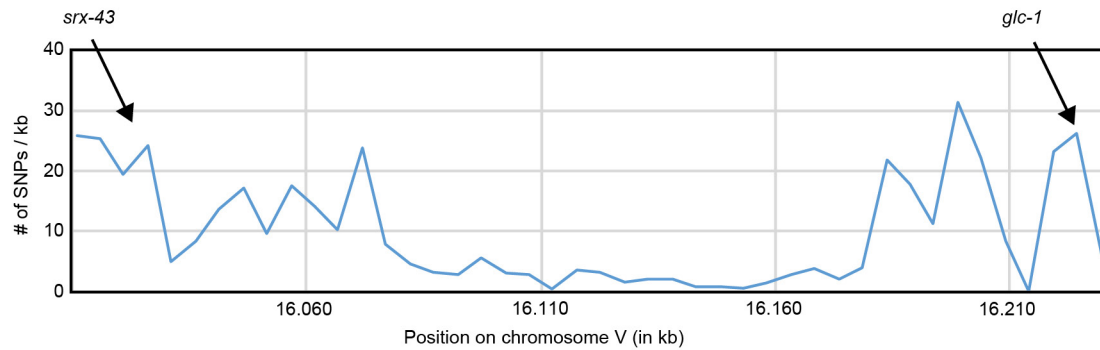




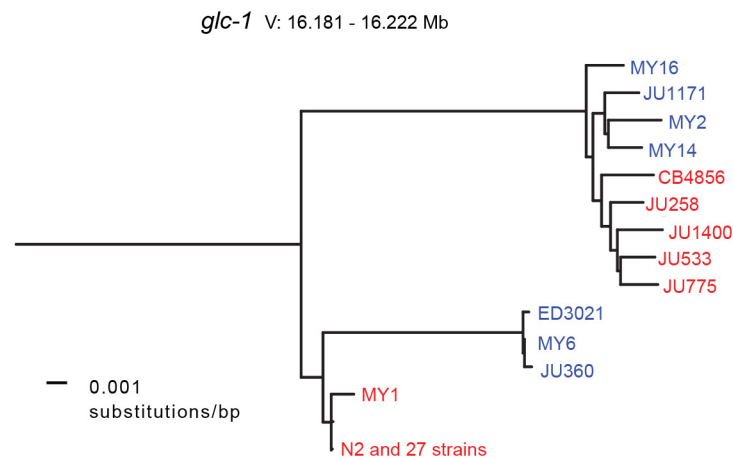
**Extended Data Figure 7 | Substantial recombination between *roam-1* and surrounding regions.** Top, phylogenies constructed with 152 diverse wild-type isolates revealing differences for the region surrounding *srx-43* and the regions immediately to the left and right of the 30-kb haplotype. Bottom, graph showing the number of variants and Tajima's *D* score

calculated for 5-kb bins across a 250-kb region. The bin containing *srx-43* has 250 polymorphisms and a Tajima's *D* of 1.01, which is high both at the genomic level (<3.4% of bins had a higher value) and for the chromosomal location of *srx-43* (<3.6% of bins had a higher value).

a

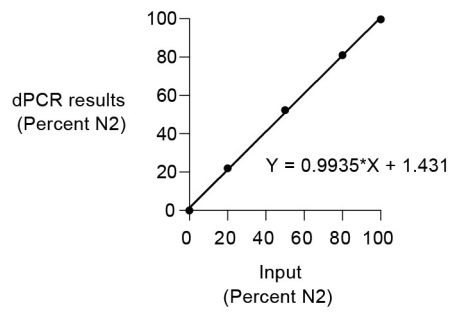


b



**Extended Data Figure 8 | Recombination between *srx-43* and *glc-1* in natural isolates.** **a**, The *glc-1* gene has previously been shown to be subject to balancing selection<sup>33</sup> and is chromosomally near *srx-43*. The blue line shows the number of SNPs per kb for N2 and MY14 averaged over 5-kb intervals for the region spanning *srx-43* and *glc-1*. The large region of low

heterozygosity between *srx-43* and *glc-1* indicates that balancing selection on *glc-1* is unlikely to account for the high heterozygosity near *srx-43*. **b**, Dendrogram for the *glc-1* region for the strains shown in Fig. 4a. The clades for *roam-1*<sub>MY14</sub> and *glc-1* are not identical.



**Extended Data Figure 9 | Standard curve for digital PCR experiments.** Best-fit line of digital PCR results for known ratios of N2 to *roam-1*<sub>MY14</sub> DNA created by mixing different ratios of genomic DNA extracted from independent N2 or *roam-1*<sub>MY14</sub> populations.



Extended Data Table 1 |  $d_N/d_S$  for *srx-43* and other genes in the *roam-1* region

Gene	Chr.	Start	Stop	substitutions	Non-synonymous	Synonymous	$dN/dS$
<i>str-231</i>	V	16006304	16007605	36	9	27	0.097
<i>str-233</i>	V	16008332	16009664	85	24	61	0.118
C06C6.1	V	16010710	16012112	72	22	50	0.128
T10C6.2	V	16012537	16014226	60	24	36	0.195
<i>str-232</i>	V	16014509	16015752	60	20	40	0.145
<i>srx-43</i>	V	16017971	16019862	30	3	27	0.033
<i>srx-44</i>	V	16020342	16022652	25	5	20	0.071
T10C6.5	V	16023351	16024185	10	1	9	0.03
T10C6.6	V	16024326	16027454	27	2	25	0.025
T10C6.7	V	16027449	16031713	12	10	2	1.26
T10C6.16	V	16032471	16034961	28	22	6	0.943
T10C6.10	V	16036413	16038293	39	22	17	0.357
<i>his-4</i>	V	16040029	16040548	6	0	6	0
<i>his-3</i>	V	16040771	16041288	13	0	13	0
<i>his-2</i>	V	16041828	16042283	3	0	3	0
<i>his-1</i>	V	16042486	16043013	7	0	7	0
T10C6.15	V	16044296	16045874	19	9	10	0.228
ZK285.2	V	16046873	16047446	8	3	5	0.175
<i>str-198</i>	V	16047808	16049147	5	3	2	0.401
<i>best-12</i>	V	16049492	16051554	13	6	7	0.227
F14H3.3	V	16051923	16053510	32	22	10	0.522
F14H3.4	V	16053900	16055072	5	3	2	0.388
F14H3.5	V	16055806	16057054	11	10	1	2.711
F14H3.6	V	16057322	16058892	11	9	2	1.24
F14H3.15	V	16059781	16060235	11	9	2	1.37
<i>fbxa-100</i>	V	16060526	16062093	11	6	5	0.303
F14H3.9	V	16067949	16068763	9	7	2	0.83
<i>cyp-35D1</i>	V	16069238	16071301	94	27	67	0.113
<i>nhr-176</i>	V	16071325	16072738	15	5	10	0.127

# Self-bound droplets of a dilute magnetic quantum liquid

Matthias Schmitt<sup>1</sup>, Matthias Wenzel<sup>1</sup>, Fabian Böttcher<sup>1</sup>, Igor Ferrier-Barbut<sup>1</sup> & Tilman Pfau<sup>1</sup>

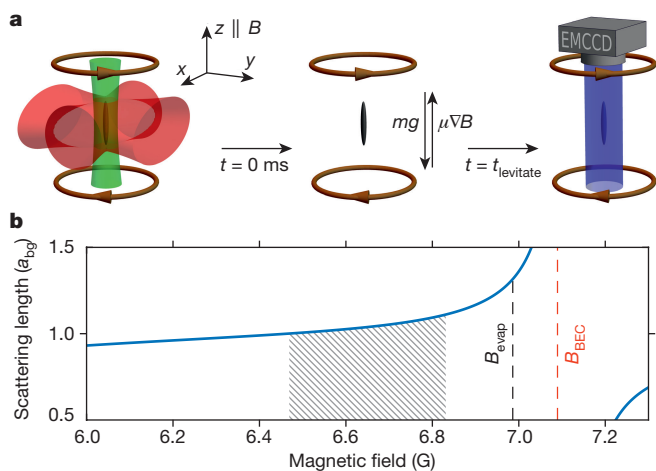
Self-bound many-body systems are formed through a balance of attractive and repulsive forces and occur in many physical scenarios. Liquid droplets are an example of a self-bound system, formed by a balance of the mutual attractive and repulsive forces that derive from different components of the inter-particle potential. It has been suggested<sup>1,2</sup> that self-bound ensembles of ultracold atoms should exist for atom number densities that are  $10^8$  times lower than in a helium droplet, which is formed from a dense quantum liquid. However, such ensembles have been elusive up to now because they require forces other than the usual zero-range contact interaction, which is either attractive or repulsive but never both. On the basis of the recent finding that an unstable bosonic dipolar gas can be stabilized by a repulsive many-body term<sup>3</sup>, it was predicted that three-dimensional self-bound quantum droplets of magnetic atoms should exist<sup>4,5</sup>. Here we report the observation of such droplets in a trap-free levitation field. We find that this dilute magnetic quantum liquid requires a minimum, critical number of atoms, below which the liquid evaporates into an expanding gas as a result of the quantum pressure of the individual constituents. Consequently, around this critical atom number we observe an interaction-driven phase transition between a gas and a self-bound liquid in the quantum degenerate regime with ultracold atoms. These droplets are the dilute counterpart of strongly correlated self-bound systems such as atomic nuclei<sup>6</sup> and helium droplets<sup>7</sup>.

Liquid droplets of water or helium are formed by the mutual attractive and repulsive forces that are created by the different parts of the inter-particle potential (and are due to covalent or van der Waals attraction and to the electronic Pauli exclusion principle, respectively). Helium droplets in particular have been a focus of research, owing to their interesting quantum nature<sup>8,9</sup>. Droplets can serve as closed, isolated quantum systems with which to probe, for example, superfluidity of mesoscopic ensembles<sup>10</sup>. In the context of ultracold atoms, the observation of an ensemble of stable droplets<sup>11</sup> in a dilute magnetic quantum gas opened up the possibility of a three-dimensional self-bound state<sup>4,5</sup>. A trapped quantum droplet of magnetic atoms has recently also been observed using erbium atoms<sup>12</sup>. Here we demonstrate the observation of dilute, self-bound liquid droplets in a sample of ultracold bosonic dysprosium atoms, which have a strong long-range magnetic dipolar interaction and a tunable repulsive short-range contact interaction. The interplay between these two interactions can be tuned such that the overall mean field is weakly attractive, but so that the interactions also create quantum depletion and a corresponding many-body repulsion. This repulsion exactly counteracts the attraction when the density of the droplet reaches the stabilization density. We use the word ‘liquid’ here to describe a state of matter that is defined by the presence of self-bound droplets and by the stabilization of the self-binding forces as a result of repulsion beyond the simple mean-field level, which manifests itself as a nontrivial correlation function. For dilute liquids, these correlations can be very weak (as in the present case), contrary to dense liquids for which correlations are strong. At small atom numbers (around 1,000 atoms), the finite size of the

wavefunction of the quantum droplet leads to a quantum pressure for each individual atom that results in an evaporation out of the self-binding potential. Therefore, these droplets are bound only above a critical atom number, which we investigate systematically.

We use <sup>164</sup>Dy, which has one of the strongest magnetic dipole moments in the periodic table with  $\mu = 9.93\mu_B$ , where  $\mu_B$  is the Bohr magneton. These atoms also offer control on the short-range interaction by a magnetic field using Feshbach resonances<sup>13–15</sup>. Here we use a specific resonance at a field of  $B_0 = 7.117(3)$  G with a width of  $\Delta B = 51(15)$  mG. (Here and elsewhere, the errors in parentheses indicate one standard deviation.) This resonance allows the scattering length  $a$  to be tuned from that of a dipole-dominated sample to a contact-dominated sample, without strong losses (Fig. 1b). To quantify the relative influence of the short-range and dipole–dipole interactions, we describe the interaction strengths using the relative dipolar strength  $\epsilon_{dd} = a_{dd}/a$ , where  $a_{dd} = \mu_0\mu^2 m / (12\pi\hbar^2) \approx 131a_0$  is the dipolar length,  $a_0$  is the Bohr radius,  $\hbar$  is the reduced Planck constant,  $\mu_0$  is the vacuum permeability and  $m$  is the atomic mass. To observe the self-bound state, we prepare an initially oblate Bose–Einstein condensate (BEC)<sup>16</sup> of <sup>164</sup>Dy with an atom number of  $N = 6,000(500)$  at a temperature of  $T = 20$  nK at large scattering length ( $B_{\text{BEC}} = 7.089(5)$  G), for which the interaction is contact-dominated, and shape it using an additional optical trap into a prolate shape along the magnetic field direction. This reshaping is done in two stages. First, we ramp up a focused beam (with a wavelength of 532 nm, aligned in the  $z$  direction) within 50 ms. With this attractive potential, the radial trap frequencies are increased to change the aspect ratio of the trap  $\lambda = \omega_z/\omega_\rho$  from  $\lambda = (80 \text{ Hz})/(20.5 \text{ Hz}) = 3.9$  down to  $\lambda = (80 \text{ Hz})/(61 \text{ Hz}) = 1.3$ ; here,  $\omega_z$  ( $\omega_\rho$ ) is the trapping frequency along (perpendicular to) the magnetic field direction. At this point, owing to magnetostriction<sup>17</sup>, the BEC becomes prolate with a cloud aspect ratio  $\kappa = \sigma_z/\sigma_\rho$  of approximately 1.5 (with  $\sigma_z$  ( $\sigma_\rho$ ) the physical size (at  $1/e^2$ ) of the cloud in (perpendicular to) the field direction) and has a typical atom number that is estimated to be  $N = 3,000(300)$ . Note that not all of these atoms are found to be in the self-bound state. Second, we apply a magnetic field gradient to the atomic cloud that exactly compensates the gravitational force and thus results in levitation. In this configuration, the cloud undergoes a continuous crossover from the BEC state directly to the single-droplet ground state as the scattering length is reduced, bypassing a bistable region<sup>4,18</sup>. Over the next 50 ms we lower the field to various values between  $B = 6.831(5)$  G and  $B = 6.469(5)$  G (indicated by the hatched area in Fig. 1b) to decrease the scattering length and create a single droplet. We hold the atoms in this configuration for 10 ms before ramping the optical trap powers within 20 ms to approximately 5% of their initial values, keeping a constant trap aspect ratio. At this point, we suddenly turn off the trap and image the cloud using far-detuned phase-contrast polarization imaging after various levitation times up to  $t_{\text{levitate}} = 90$  ms. This sequence is shown schematically in Fig. 1a. Being sensitive only to high densities, we observe that a thermal fraction of the atomic cloud expands very quickly, whereas a very small and dense cloud remains for very long times. We interpret this observation

<sup>1</sup>Physikalisches Institut and Center for Integrated Quantum Science and Technology, Universität Stuttgart, Pfaffenwaldring 57, 70550 Stuttgart, Germany.

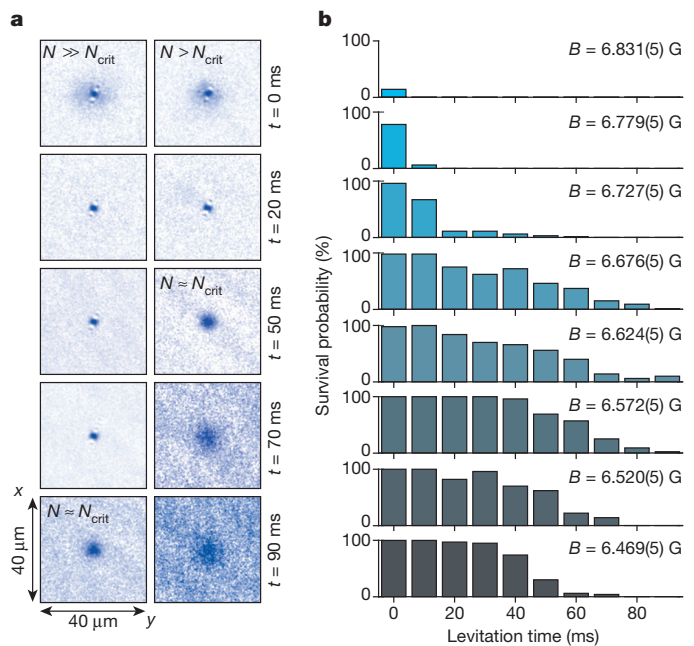


**Figure 1 | Experimental sequence.** **a**, A schematic of the experimental sequence. We start with an atomic ensemble in a crossed optical trap superimposed with a magnetic field gradient that is strong enough to compensate the gravitational force. We then turn off the trapping beams and levitate the droplet for various times  $t_{\text{levitate}}$ . Finally, we image the atoms using phase-contrast polarization imaging projected on an EMCCD (electron multiplying charge coupled device) camera. **b**, Scattering length as a function of magnetic field at the region of the Feshbach resonance in units of the positive background scattering length  $a_{\text{bg}}$ . The red dashed line indicates the field ( $B_{\text{BEC}}$ ) at which we create a BEC. The hatched area describes the region in which the experiments were performed. The dashed black line shows the magnetic field ( $B_{\text{evap}}$ ) used to intentionally evaporate the droplets to the gas phase.

as a self-bound quantum droplet. We calculated the radial size of the quantum droplets to be approximately 300 nm, which is smaller than our imaging resolution of  $1 \mu\text{m}$  such that we observe astigmatic diffraction (see Fig. 2a). At specific fields, we observe these droplets for times as long as  $t_{\text{levitate}} = 90$  ms. At some time during the trap-free levitation, we observe that the droplets have expanded. We reason that this behaviour is due to the fact that droplets lose atoms as a result of three-body decay or residual excitations until they reach a critical atom number, below which they are no longer self-bound and evaporate back into a gas phase. In this context, the word ‘evaporation’ is used to denote the transition from a dilute self-bound liquid state to an expanding gas state. Given our shot-to-shot noise in the initial atom number, the critical atom number is reached at different times. This behaviour is represented in Fig. 2a.

As a first analysis, we count the images in which we observe a single droplet over 100 realizations and plot the survival probability for different magnetic fields (Fig. 2b). The levitation time is varied between  $t_{\text{levitate}} = 0$  ms, which essentially represents a trapped cloud, and  $t_{\text{levitate}} = 90$  ms. For low scattering length ( $B = 6.469(5)$  G), we always create a single droplet, but its lifetime is short. As the scattering length increases, so does the lifetime. We find a maximal survival probability in the magnetic field range  $B = 6.572(5)$ – $6.676(5)$  G. For even higher scattering lengths, we find droplets only at 0 ms, and very few self-bound droplets. The calculated survival probabilities are in qualitative agreement with an increasing critical atom number and a decreasing rate of atom loss in the droplets with increasing scattering length. This behaviour has been observed<sup>3</sup> in a waveguide configuration and for a single trapped droplet<sup>12</sup>, and is supported by calculations on a self-bound droplet<sup>5,19</sup>. However, the precise evolution depends on the spread in initial atom number and the fact that droplets evaporate at different atom numbers (see below).

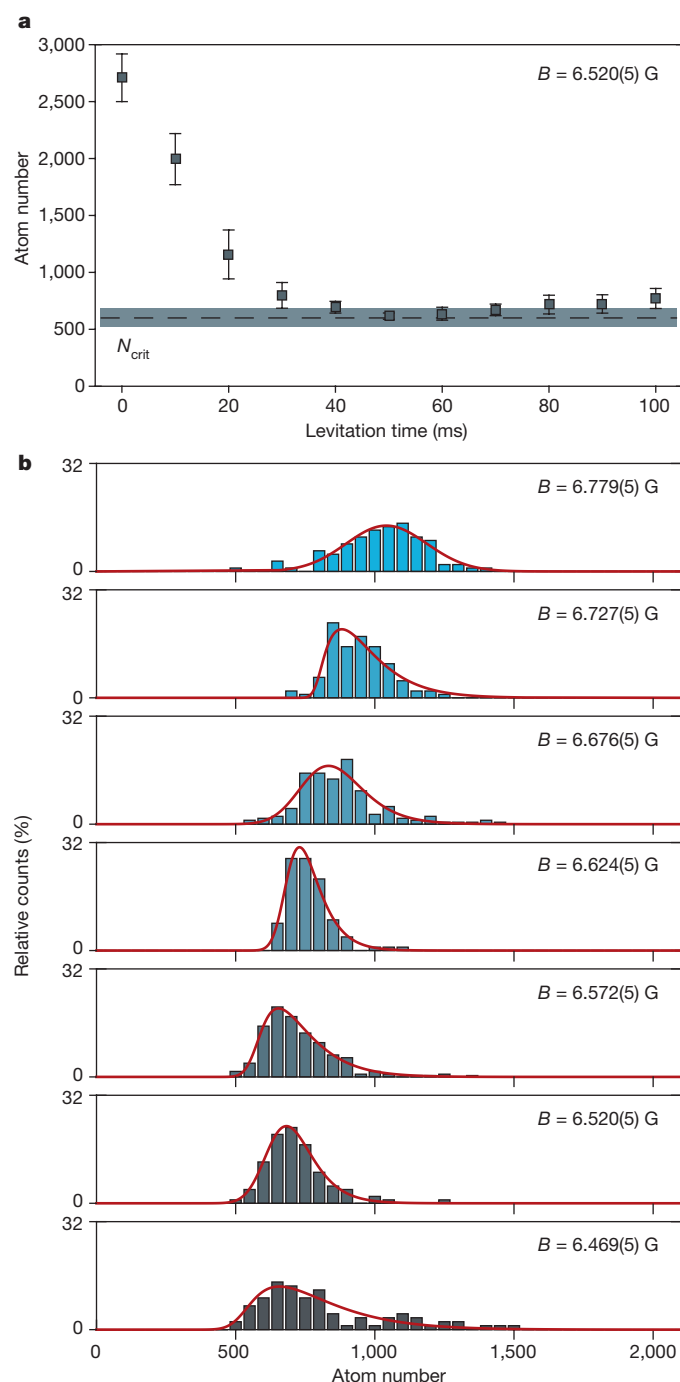
To obtain a more quantitative analysis of the critical atom number of these droplets, we intentionally evaporate them after variable levitation times by increasing the magnetic field to  $B_{\text{evap}} = 6.986(5)$  G (dashed black line in Fig. 1b). At this field, we observe that all droplets have been evaporated and interpret this to mean that the critical atom number at



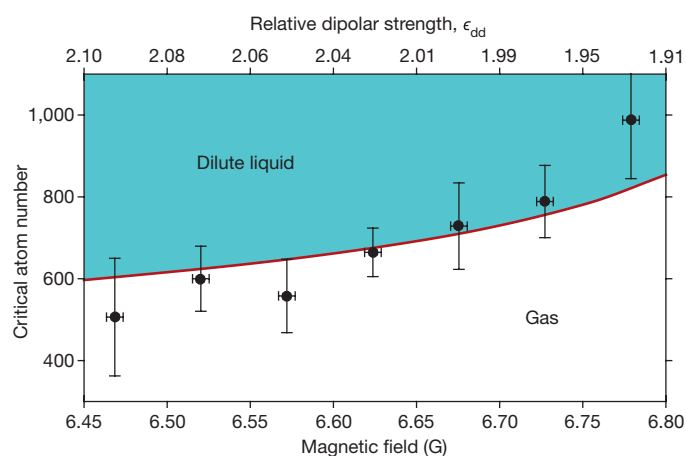
**Figure 2 | Droplet survival probability.** **a**, Image sequences of two droplets with different levitation times at the same magnetic field  $B = 6.676(5)$  G. The images are not multiple images of the same droplet, but are selected from various images because the imaging process is destructive. All images are rescaled to the maximum optical density and have been re-centred. In the left-hand column, we start with an atom number that is much larger than the critical atom number for stable droplets and observe a single droplet up to  $t_{\text{levitate}} = 70$  ms. Between  $t = 70$  ms and  $t = 90$  ms, the cloud reaches the critical atom number and evaporates back to a gas phase, observed as an expanding cloud. In the right-hand column, the droplet starts with an atom number that is much closer to the critical atom number, leading to an earlier evaporation, between 20 ms and 50 ms of levitation time. From this point, the cloud evaporates to the BEC phase and expands. **b**, Histogram of the survival probability of a single droplet as function of levitation time and magnetic field. At low scattering lengths ( $B = 6.469(5)$  G), we always observe droplets for up to  $t_{\text{levitate}} = 30$  ms, followed by a fast decay in survival probability that is explained by a fast decay in atom number as a result of three-body collisions. For increasing scattering length, we observe an increase in the lifetime of the droplets up to a magnetic field range of  $B = 6.572(5)$ – $6.676(5)$  G. At these conditions we observe a single droplet with a size below our resolution after a levitation time of  $t_{\text{levitate}} = 90$  ms. Further increase of the scattering length leads to a fast decay of self-bound droplets even for short times ( $t_{\text{levitate}} = 20$  ms), which we interpret as originating from an increase in the critical atom number to values close to our initial atom number. For the highest scattering length ( $B = 6.831(5)$  G), we barely create droplets in the trap.

this field is higher than all relevant atom numbers observed here. After expansion, the atom number can be determined accurately without being limited by the finite resolution of the imaging optics. Here we observe that the number of atoms in the droplets decays to an essentially constant number—further indication of a critical atom number for self-bound droplets. This behaviour is demonstrated in Fig. 3a for a magnetic field of  $B = 6.520(5)$  G, wherein each point is represented by a mean atom number that is calculated from 20 images, and the error denotes one standard deviation. A histogram of the atom number distribution for long levitation times ( $t_{\text{levitate}} \geq 60$  ms) and for different magnetic fields is shown in Fig. 3b. We observe that the atom number distributions shift with scattering length, and conclude that the droplets lose atoms until they reach the critical atom number, at which point all of the atoms evaporate out of the droplets into the gas phase. We observe that at long times, when most droplets have evaporated, there is an asymmetric dispersion in atom number to higher values. We posit that this reflects the fact that not all droplets evaporate at exactly the





**Figure 3 | Critical atom number.** **a**, Decay in atom number as function of levitation time. Each point represents the mean atom number of 20 realizations; error bars denote the standard deviation. We observe a decay for short times to an essentially constant number for long times. The dashed line shows the critical atom number as determined by the best fit of our convoluted model (see text) to the data; the shaded area shows the error in this fit, represented by the quadratic mean of the widths of the convoluted distributions. For levitation times of less than 20 ms, the measured atom number overestimates the atom number in the self-bound droplet because it is hard to distinguish the droplet from the surrounding BEC in our experimental procedure. **b**, We analyse the atom number distribution for levitation times in the range  $t_{\text{levitate}} = 60\text{--}100$  ms because the atom number is mostly constant in this range. We bin the atom number to a window of 50 atoms and plot the relative counts as a function of atom number and magnetic field. The red curves represent fits of the convoluted functions to the observed histograms. The colours of the plotted histograms match those in Fig. 2b, and represent the magnetic field.



**Figure 4 | Phase transition between dilute liquid and gas.** The data points show the critical atom number as a function of the magnetic field, as determined from the fit values from Fig. 3b. The error in the atom number is given by the quadratic mean of the widths of the Gaussian and Maxwell–Boltzmann distributions; the error in magnetic field describes the resolution of our magnetic field coils. As the magnetic field decreases, so does the critical atom number  $N_{\text{crit}}$ . We identify the upper-left corner as the dilute liquid phase and the lower-right corner as the gas phase. The critical atom number describes the phase transition between dilute liquid and gas. The solid red line represents a full Gross–Pitaevskii simulation for different relative dipolar strengths  $\epsilon_{\text{dd}}$ .

critical atom number, but that the evaporation for some droplets can occur with  $N > N_{\text{crit}}$  owing to the presence of residual (for example, thermal) excitations in the droplet. To extract a critical atom number we fit the histograms with a convolution of a Gaussian and a Maxwell–Boltzmann distribution (see Methods).

The best fits are shown in Fig. 3b as red curves. We plot our result of the critical atom number in Fig. 4 and compare it to full, extended Gross–Pitaevskii simulations (see Methods). The error is given by the quadratic mean of the widths of the Maxwell–Boltzmann and Gaussian distributions. This way of determining  $N_{\text{crit}}$  depends on the model used to determine the fit; other definitions could lead to slightly different values. Nevertheless, we see a clear change in the critical atom number with magnetic field, and with this we probe the phase transition line between the dilute liquid phase and the gas phase. To compare the results with the simulations, we calculate the relative dipolar strength for our magnetic field range. To do so, we include the Feshbach resonance at  $B_{01} = 7.117(3)$  G with a width of  $\Delta B_1 = 51(15)$  mG and a resonance at  $B_{02} = 5.1(1)$  G with a width of  $\Delta B_2 = 0.1(1)$  G. A best fit is obtained when we change the previously assumed local background scattering length<sup>14</sup> of  $a_{\text{bg}} = 92(8)a_0$  to  $a_{\text{bg}} = 62.5a_0$ . This lower value seems, at first, to be incompatible with previous measurements at different fields<sup>14</sup>; however, the complexity of the scattering problem in dysprosium does not allow a theoretical prediction and the local  $a_{\text{bg}}$  might vary in other ranges of magnetic field. In addition, theoretical simulations of the Rosensweig instability<sup>18,19</sup> suggest that a background scattering length of less than  $92a_0$  is necessary to agree with experimentally observed timescales<sup>11</sup>. In our measurements, the strong dependence of  $N_{\text{crit}}$  on scattering length provides a very high sensitivity. Changing the background scattering length from  $92a_0$  to  $62.5a_0$  reduces  $N_{\text{crit}}$  by almost a factor of ten. This method therefore enables a very precise measurement of the scattering length. However, at this level of precision, we must question the approximations made in our model, such as the first-order Born approximation for the dipolar scattering and the local density approximation. Consequently, the value of  $a_{\text{bg}}$  quoted here is model-dependent, and could be subject to future corrections. An independent measurement of  $a$ , via the methods of ref. 12 for instance, would make  $N_{\text{crit}}$  measurements a very sensitive benchmark for many-body theories.

By removing the need for any trapping potential, our observation of the self-bound regime offers access to truly isolated, dissipative quantum systems in which the effective cancellation of the mean field enables quantum correlations to be studied in detail. The gas-to-liquid transition and, in particular, the nucleation dynamics of the droplets will be sensitive probes of the interplay between interactions and quantum correlations.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 25 July; accepted 29 September 2016.**

- Petrov, D. S. Quantum mechanical stabilization of a collapsing Bose–Bose mixture. *Phys. Rev. Lett.* **115**, 155302 (2015).
- Bulgac, A. Dilute quantum droplets. *Phys. Rev. Lett.* **89**, 050402 (2002).
- Ferrier-Barbut, I., Kadau, H., Schmitt, M., Wenzel, M. & Pfau, T. Observation of quantum droplets in a strongly dipolar Bose gas. *Phys. Rev. Lett.* **116**, 215301 (2016).
- Wächtler, F. & Santos, L. Ground-state properties and elementary excitations of quantum droplets in dipolar Bose–Einstein condensates. *Phys. Rev. A* **94**, 043618 (2016).
- Baillie, D., Wilson, R. M., Bisset, R. N. & Blakie, P. B. Self-bound dipolar droplet: a localized matter wave in free space. *Phys. Rev. A* **94**, 021602(R) (2016).
- Bender, M., Heenen, P.-H. & Reinhard, P.-G. Self-consistent mean-field models for nuclear structure. *Rev. Mod. Phys.* **75**, 121–180 (2003).
- Dalfovo, F. & Stringari, S. Helium nanodroplets and trapped Bose–Einstein condensates as prototypes of finite quantum fluids. *J. Chem. Phys.* **115**, 10078–10089 (2001).
- Volovik, G. *The Universe in a Helium Droplet* 27–31 (Oxford Univ. Press, 2009).
- Toennies, J. P. & Vilesov, A. F. Superfluid helium droplets: a uniquely cold nanomatrix for molecules and molecular complexes. *Angew. Chem. Int. Ed.* **43**, 2622–2648 (2004).
- Gomez, L. F. *et al.* Shapes and vorticities of superfluid helium nanodroplets. *Science* **345**, 906–909 (2014).
- Kadau, H. *et al.* Observing the Rosensweig instability of a quantum ferrofluid. *Nature* **530**, 194–197 (2016).
- Chomaz, L. *et al.* Quantum-fluctuation-driven crossover from a dilute Bose–Einstein condensate to a macro-droplet in a dipolar quantum fluid. Preprint at <https://arxiv.org/abs/1607.06613> (2016).
- Chin, C., Grimm, R., Julienne, P. & Tiesinga, E. Feshbach resonances in ultracold gases. *Rev. Mod. Phys.* **82**, 1225–1286 (2010).
- Tang, Y., Sykes, A., Burdick, N. Q., Bohn, J. L. & Lev, B. L. s-wave scattering lengths of the strongly dipolar bosons  $^{162}\text{Dy}$  and  $^{164}\text{Dy}$ . *Phys. Rev. A* **92**, 022703 (2015).
- Maier, T. *et al.* Emergence of chaotic scattering in ultracold Er and Dy. *Phys. Rev. X* **5**, 041029 (2015).
- Lu, M., Burdick, N. Q., Youn, S. H. & Lev, B. L. Strongly dipolar Bose–Einstein condensate of dysprosium. *Phys. Rev. Lett.* **107**, 190401 (2011).
- Lahaye, T., Menotti, C., Santos, L., Lewenstein, M. & Pfau, T. The physics of dipolar bosonic quantum gases. *Rep. Prog. Phys.* **72**, 126401 (2009).
- Bisset, R. N., Wilson, R. M., Baillie, D. & Blakie, P. B. Ground-state phase diagram of a dipolar condensate with quantum fluctuations. *Phys. Rev. A* **94**, 033619 (2016).
- Wächtler, F. & Santos, L. Quantum filaments in dipolar Bose–Einstein condensates. *Phys. Rev. A* **93**, 061603 (2016).

**Acknowledgements** We thank H. P. Büchler, L. Santos, F. Ferlaino, W. Ketterle, H. Sadeghpour, M. Zwierlein and V. Vuletić for discussions. This work is supported by the German Research Foundation (DFG) within SFB/TRR21 as well as FOR 2247. I.F.-B. acknowledges support from the EU within Horizon2020 Marie Skłodowska Curie IF (703419 DiplnQuantum).

**Author Contributions** All authors discussed the results, made critical contributions to the work and contributed to the writing of the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.S. ([m.schmitt@physik.uni-stuttgart.de](mailto:m.schmitt@physik.uni-stuttgart.de)) or T.P. ([tpfau@physik.uni-stuttgart.de](mailto:tpfau@physik.uni-stuttgart.de)).

**Reviewer Information** *Nature* thanks B. Blakie, R. Hulet and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

**Convolution model.** To extract the critical atom number from the data in Fig. 3 we fit the histograms with a phenomenological model (represented as red lines). This model consists of the convolution of a Gaussian and a Maxwell–Boltzmann distribution. The symmetric Gaussian distribution represents broadening effects that result from statistical errors including detection noise. The asymmetric Maxwell–Boltzmann distribution is used to model the possibility of a droplet fully evaporating at atom numbers higher than the critical atom number, as a result of the presence of collective excitations in the droplets. From the fit we extract the critical atom number and two widths, one from each distribution in the convolution. We represent the quadratic mean of these widths as error bars in Fig. 4.

**Extended Gross–Pitaevskii simulation.** To compare our results to current theory<sup>4,5</sup>, we perform simulations of the effective Gross–Pitaevskii equation

$$i\hbar\partial_t\psi = \left[ -\frac{\hbar^2\nabla^2}{2m} + V_{\text{ext}} + g|\psi|^2 + \int V_{\text{dd}}(\mathbf{r}-\mathbf{r}')|\psi(\mathbf{r}')|^2 d\mathbf{r}' + \frac{32g\sqrt{a^3}}{3\sqrt{\pi}}\left(1 + \frac{3}{2}\epsilon_{\text{dd}}^2\right)|\psi|^3 - i\frac{\hbar}{2}L_3|\psi|^4 \right]\psi(\mathbf{r}) \quad (1)$$

using a simple interaction potential, and taking into account quantum fluctuations within a local density approximation<sup>20,21</sup> and three-body losses. Here

$$V_{\text{dd}}(\mathbf{r}) = \frac{\mu_0\mu^2}{4\pi} \frac{1 - 3\cos^2(\vartheta)}{|\mathbf{r}|^3}$$

describes the dipole–dipole interaction potential, with  $\vartheta$  denoting the angle between the polarization direction of the dipoles and their relative orientation.

The main assumptions of this model are therefore the validity of the local density approximation and of the interaction potential, which results from the first-order Born approximation. The magnetic moment and scattering length are  $\mu = 9.93\mu_B$  and  $a = 60a_0 - 80a_0$ , respectively. The latter defines  $g = 4\pi a\hbar^2/m$  and is chosen such that we are in agreement with the critical atom numbers we observe in the experiment. The loss parameter  $L_3 = 1.25 \times 10^{-41} \text{ m}^6 \text{ s}^{-1}$  is estimated from measurements on a thermal cloud and is assumed to be constant over the small range of scattering lengths. The validity of the local density approximation is supported by quantum Monte Carlo simulations<sup>22</sup> and recent measurements with erbium atoms<sup>12</sup>.

To obtain the data in Fig. 4, we choose  $V_{\text{ext}} = 0$  and initially prepare  $N_0 > N_{\text{crit}}$  atoms with a Gaussian density distribution ( $\sigma_r = 250 \text{ nm}$ ,  $\sigma_z = 1,500 \text{ nm}$ ). The ground state is reached by imaginary time evolution of equation (1) using a split-step Fourier method. Following this preparation of the self-bound droplet with  $N_0$  atoms, we simulate the dynamics via real-time evolution. Because the atom number  $N < N_0$  decays, owing to three-body losses, the density and the effective two-body attraction are also reduced. At  $N = N_{\text{crit}}$ , the contributions by the effective two-body attraction and the quantum pressure are the same in magnitude, and the droplet evaporates quickly. This evaporation process manifests itself as a decrease in peak density of at least one order of magnitude. Three-body losses are highly suppressed then, such that the atom number stays almost constant for an evaporated droplet.

**Data availability.** The data that support the findings of this study are available from the corresponding author upon reasonable request.

20. Lima, A. R. P. & Pelster, A. Quantum fluctuations in dipolar Bose gases. *Phys. Rev. A* **84**, 041604 (2011).
21. Lima, A. R. P. & Pelster, A. Beyond mean-field low-lying excitations of dipolar Bose gases. *Phys. Rev. A* **86**, 063609 (2012).
22. Saito, H. Path-integral Monte Carlo study on a droplet of a dipolar Bose–Einstein condensate stabilized by quantum fluctuation. *J. Phys. Soc. Jpn* **85**, 053001 (2016).



# Tracking the ultrafast motion of a single molecule by femtosecond orbital imaging

Tyler L. Cocker<sup>1\*</sup>, Dominik Peller<sup>1\*</sup>, Ping Yu<sup>1</sup>, Jascha Repp<sup>1</sup> & Rupert Huber<sup>1</sup>

Watching a single molecule move on its intrinsic timescale has been one of the central goals of modern nanoscience, and calls for measurements that combine ultrafast temporal resolution<sup>1–8</sup> with atomic spatial resolution<sup>9–30</sup>. Steady-state experiments access the requisite spatial scales, as illustrated by direct imaging of individual molecular orbitals using scanning tunnelling microscopy<sup>9–11</sup> or the acquisition of tip-enhanced Raman and luminescence spectra with sub-molecular resolution<sup>26–28</sup>. But tracking the intrinsic dynamics of a single molecule directly in the time domain faces the challenge that interactions with the molecule must be confined to a femtosecond time window. For individual nanoparticles, such ultrafast temporal confinement has been demonstrated<sup>18</sup> by combining scanning tunnelling microscopy with so-called lightwave electronics<sup>1–8</sup>, which uses the oscillating carrier wave of tailored light pulses to directly manipulate electronic motion on timescales faster even than a single cycle of light. Here we build on ultrafast terahertz scanning tunnelling microscopy to access a state-selective tunnelling regime, where the peak of a terahertz electric-field waveform transiently opens an otherwise forbidden tunnelling channel through a single molecular state. It thereby removes a single electron from an individual pentacene molecule's highest occupied molecular orbital within a time window shorter than one oscillation cycle of the terahertz wave. We exploit this effect to record approximately 100-femtosecond snapshot images of the orbital structure with sub-ångström spatial resolution, and to reveal, through pump/probe measurements, coherent molecular vibrations at terahertz frequencies directly in the time domain. We anticipate that the combination of lightwave electronics<sup>1–8</sup> and the atomic resolution of our approach will open the door to visualizing ultrafast photochemistry and the operation of molecular electronics on the single-orbital scale.

We investigate the  $\pi$ -conjugated organic molecule pentacene with low-current<sup>9</sup>, low-temperature terahertz scanning tunnelling microscopy (THz-STM). In our experiments, we focus phase-stable, single-cycle THz pulses<sup>7,8</sup> onto the STM junction. There, the electric field of the THz waveform adds to the static electric potential between tip and sample, resulting in an ultrafast modulation of the applied bias voltage. To preserve the inherent electronic properties of pentacene, the molecules are adsorbed on monolayer islands of NaCl, which electronically decouple the pentacene molecules from the underlying Au(110) surface<sup>9</sup> (Fig. 1a; see Methods for further details of sample preparation and microscope set-up).

Figure 1b depicts the novel ultrafast, energy-selective tunnelling regime at the heart of this study. It builds on the orbital-specific tunnelling of conventional STM that is enabled by the use of insulating films<sup>9</sup>, where the double-barrier (salt, vacuum) geometry leads to two very distinct tunnelling mechanisms (Fig. 1b, left). For low bias voltages, tunnelling electrons cannot access the molecular orbitals and must tunnel between the tip and the metal substrate in a single step. Conversely, once the voltage reaches a threshold such that either the

lowest unoccupied molecular orbital (LUMO) or the highest occupied molecular orbital (HOMO) is inside the bias voltage window, sequential tunnelling sets in. In this two-step process, for example, an electron will temporarily localize in the LUMO after having tunneled through the first barrier, and before tunnelling through the second barrier. The alignment of the LUMO or HOMO transport level with the Fermi energy of the tip thus opens up a new tunnelling channel, and consequently appears as a peak in the differential conductance (Fig. 1b, centre).

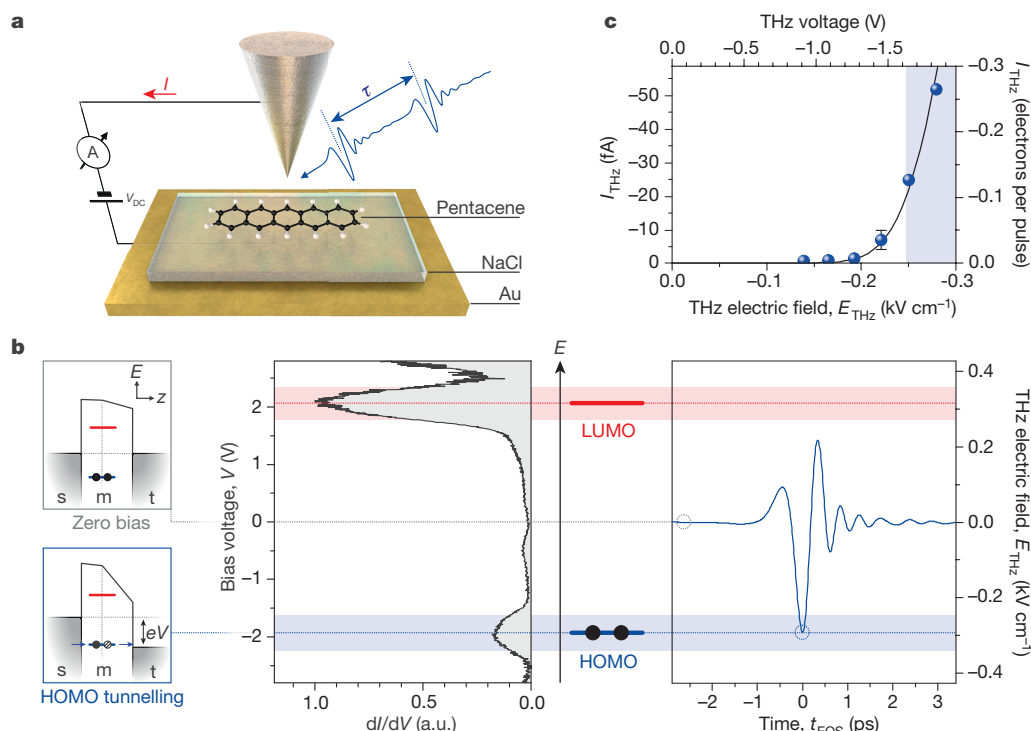
To understand how this situation enables a new regime in THz-STM, we now consider the dynamic analogue of this process, where a THz pulse replaces the steady-state bias (Fig. 1b, right). Light-matter interactions with intense THz fields are often characterized by different competing quantum effects<sup>4–6</sup>, occurring at the crossover of the photon versus classical-field regimes. The semi-classical picture discussed in the following is justified by the experimental results presented further below, but cannot be assumed *a priori*. To facilitate the discussion, however, we will introduce the new regime directly in the context of this picture, in which the THz waveform is viewed as a transient bias voltage and hence a transient modification to the level alignment of the system. Thus, it is expected to temporarily open an otherwise forbidden tunnelling channel through the HOMO when the peak of the waveform matches the corresponding transport level.

This concept of THz-induced sequential tunnelling has the potential to yield optimally short current responses that are ultimately limited only by the linewidth of the resonance. Moreover, the asymmetry of the THz waveform should allow us to drive tunnelling out of the HOMO without accessing the LUMO, and create a unipolar current burst with a duration defined by the width of the THz waveform crest. Finally, we note that sensitivity to rectified currents below one electron per THz pulse is a prerequisite for exploring this new ultrafast, state-selective tunnelling regime. In the experiment, we position the STM tip over a pentacene molecule, turn off the STM feedback loop as well as the bias voltage, and measure the net THz-induced tunnel current through the molecule as a function of the THz peak field. We observe a rectified current onset in the negative bias direction (Fig. 1c) and currents approaching one electron per THz pulse. The threshold peak field is about  $0.25 \text{ kV cm}^{-1}$ , corresponding to a peak bias voltage of about  $-1.65 \text{ V}$ , owing to the large field enhancement<sup>6–8,18,30</sup> at the tip.

The ultimate test of whether this signal is due to THz-induced sequential tunnelling is the spatial distribution of the rectified current. Steady-state STM imaging with the bias set to the HOMO transport level directly reveals the HOMO density contours of the free molecule<sup>9</sup>. Figure 2a shows such a steady-state, constant-current STM image of a pentacene molecule, recorded at a voltage near the HOMO level (bias voltage,  $V_{\text{DC}} = -1.7 \text{ V}$ ), and features the ten lobes of the HOMO protruding towards the tip<sup>9</sup>. In contrast, these lobes are absent from the steady-state STM image when the bias voltage window does not include a molecular resonance (shown in Fig. 2b,  $V_{\text{DC}} = +6 \text{ mV}$ ) and only one-step tunnelling is permitted. The THz-STM image recorded

<sup>1</sup>Department of Physics, University of Regensburg, 93040 Regensburg, Germany.

\*These authors contributed equally to this work.



**Figure 1 | Ultrafast THz-induced tunnelling of a single electron out of a single molecule.** **a**, Our experimental set-up. Two THz pulses separated by a variable delay time,  $\tau$ , are coupled into the STM junction, which includes a tungsten tip and a pentacene molecule that is electronically decoupled from the Au(110) substrate by a monolayer NaCl film. The net current produced by the THz pulses was measured as a shift in the DC tunnel current,  $I$ . The bias voltage,  $V_{DC}$ , refers to the voltage of the metal substrate relative to the tip. Likewise, positive currents correspond to electrons tunnelling from the tip into the sample. **b**, Concept of steady-state and ultrafast tunnelling out of the HOMO of a single molecule. For low bias voltages (far left, top), tunnelling electrons cannot access the molecular resonances. In contrast, for voltages exceeding the LUMO or HOMO transport levels (far left, bottom), sequential tunnelling sets in.  $E$ , energy;  $e$ , elementary charge;  $m$ , pentacene molecule;  $s$ , Au substrate;  $t$ , tip;  $V$ , applied voltage;  $z$ , position along substrate–tip axis. Sequential tunnelling (centre) results in peaks in the steady-state differential

conductance ( $dI/dV$ ). In the dynamic case (right), the THz pulse acts as an ultrafast voltage transient. Here, the steady-state bias is  $V_{DC} = 0$  and the THz peak field reaches the HOMO but not the LUMO, giving rise to state-selective sequential tunnelling through the HOMO. The waveform (blue curve) is the electro-optically measured far-field trace used in the experiment. **c**, THz-induced tunnel current ( $I_{THz}$ ) out of a single pentacene molecule on monolayer NaCl as a function of THz field strength ( $E_{THz}$ ), showing the onset of electron tunnelling out of the HOMO ( $V_{DC} = 0$  V). Rescaling the time-average current (left axis) to electrons tunnelled per THz pulse (right axis; see Methods) reveals currents of less than 1 electron per pulse. Under the assumption that the THz electric field results in a voltage transient with a scaling factor determined by the tip-enhancement of the field (see Extended Data Fig. 3), the onset of the THz-induced current is well reproduced (solid line). The error bar indicates the standard deviation from the fit.

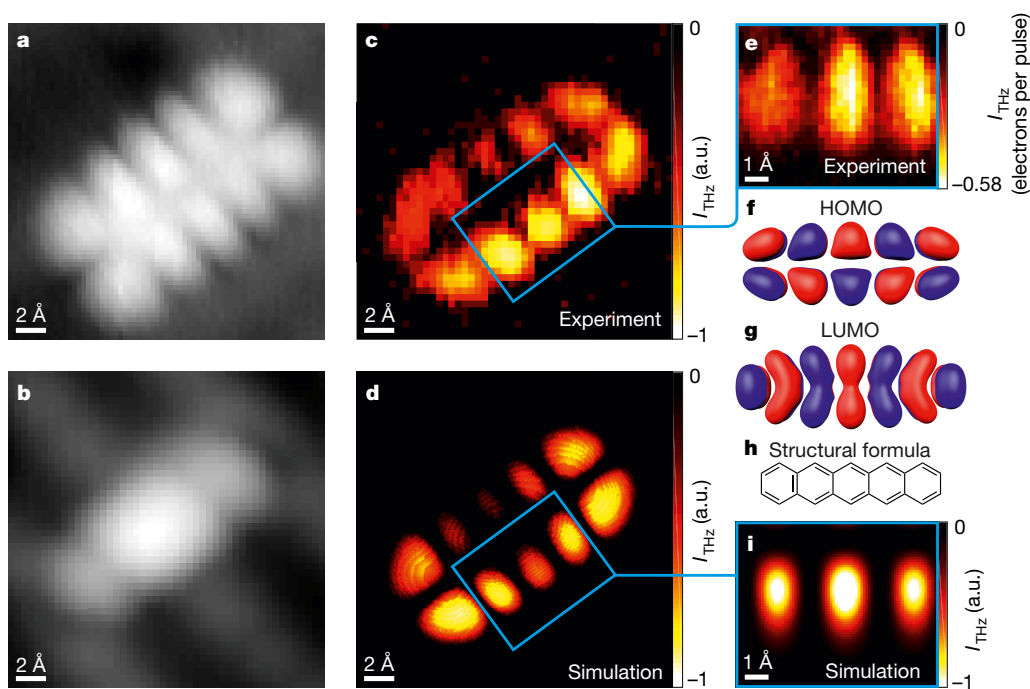
simultaneously, however, differs dramatically (Fig. 2c). It clearly shows ten lobes in the spatial distribution of the rectified current, and is strikingly similar to the HOMO density, indicating that the THz waveform indeed drives tunnelling through the HOMO. The contrast between the lobes of the HOMO even appears sharper in the THz-STM image than in the conventional image Fig. 2a.

For easy comparison, Fig. 2f and Fig. 2g show, respectively, the HOMO and LUMO determined from density functional theory (DFT) calculations, alongside the structural formula of pentacene (Fig. 2h). The central nodal plane of the HOMO, which is not present in the LUMO, can be clearly identified in the THz-STM images. The deviations of the THz-STM image (Fig. 2c) from the HOMO density are explained by the fact that the tip follows the low-voltage topography during image acquisition. This can be seen in the corresponding image simulation (Fig. 2d), in which we display the square of the matrix element between the tip wavefunction and the HOMO for tip positions following the experimental low-voltage topography acquired simultaneously (Fig. 2b). The simulation reproduces the THz-STM image in great detail (compare Fig. 2d and Fig. 2c), including the asymmetry in the orbital, which arises from a slight shift of the low-bias topography relative to the HOMO, probably due to an asymmetric tip shape.

Ultrafast THz-STM imaging can also be performed with the STM bias voltage and feedback loop completely turned off. This mode avoids artefacts from the low-bias topography. Figure 2e shows

a constant-height THz-STM scan ( $V_{DC} = 0$  V) of three central lobes. Here, the THz pulses apply the only voltage to the system. Nevertheless, with their peak voltages set near the HOMO resonance, they induce a rectified current on the scale of 0.6 electrons per THz pulse, revealing orbital lobes (simulated in Fig. 2i) with detail on the scale of 0.6 Å (see Extended Data Fig. 1 and Methods). Hence, not only can a THz pulse be used to remove a single electron from a single molecule with femtosecond control over the tunnelling time window, but this can also be done with sub-ångström precision and orbital selectivity. As the THz pulses are phase stable, this tunnelling time window is locked to a fixed point within the pulse envelope irrespective of the tip position, resulting in a stroboscopic femtosecond snapshot image of a single orbital recorded pixel by pixel. Under different experimental conditions, we can also trace the LUMO with the THz-driven tunnel current (see Extended Data Fig. 2).

The time resolution of the snapshot in Fig. 2 can be extracted from a THz autocorrelation<sup>18</sup> measurement over a central pentacene HOMO lobe (data points in Fig. 3a), where the rectified current is recorded as a function of the delay,  $\tau$ , between two identical THz pulses (Fig. 3a, top). To simulate the experimental trace, we account for the tunnelling conductance of the HOMO and LUMO levels and compute the rectified current induced by time-delayed pairs of THz waveforms, as measured by electro-optic sampling (see Methods for details), under the assumption that the voltage transient equals the THz electric field



**Figure 2 | Ultrafast THz-STM imaging of a pentacene molecular orbital.** **a**, Steady-state, constant-current STM image of the HOMO at  $V_{\text{DC}} = -1.7$  V and  $I = 0.83$  pA. Greyscale range =  $2.3 \text{ \AA}$ . **b**, Steady-state STM image at  $V_{\text{DC}} = 6$  mV and  $I = 2.2$  pA. The Au(110) missing-row reconstruction beneath the NaCl film gives rise to an overall stripe pattern that affects the low-voltage appearance of pentacene, which is otherwise featureless<sup>9</sup>. Greyscale range =  $2.0 \text{ \AA}$ . **c**, THz-STM image taken at the same time as the topography image in **b**, with the THz voltage set to  $V_{\text{THz}}^{\text{peak}} \approx -2.05$  V. The THz-induced current was in the direction of electron tunnelling from the HOMO to the tip, even though the DC bias was

positive and far from either molecular resonance. **d**, Simulated THz-STM image calculated on the basis of the Tersoff–Hamann approximation, with the tip position following the topography shown in **b**. **e**, THz-STM image (close-up of the blue rectangle in **c**) measured at constant height, with  $V_{\text{DC}} = 0$  V,  $V_{\text{THz}}^{\text{peak}} \approx -2.05$  V, and a time resolution of 115 fs. The THz-induced current is calibrated in units of rectified electrons per THz pulse. **f**, **g**, DFT-derived HOMO (**f**) and LUMO (**g**) contours of the free pentacene molecule. **h**, Molecular structure of pentacene. **i**, Simulated constant-height THz-STM image (close-up of blue rectangle in **d**).

multiplied by a scaling factor, which acts as the only fit parameter. Good agreement is found between the simulated (Fig. 3a, black curve) and measured (Fig. 3a, data points) autocorrelations for a peak voltage of  $-2.22 \text{ V} \pm 0.07 \text{ V}$  at  $\tau = 0$  fs, supporting our assertion that THz pulses tuned to the HOMO resonance drive sequential tunnelling. Yet more fundamentally, this result verifies that the THz waveform modulates the bias voltage of the junction quasi-instantaneously, even on the sub-cycle scale. In this regime, the simulations allow us to retrieve the average temporal shape of the rectified current pulse. As shown in Extended Data Fig. 3, the rectified current is generated within a time window as short as 130 fs for a voltage peak of  $-2.22$  V. Furthermore, we estimate that the THz peak was only  $-2.05$  V during THz-STM imaging (Fig. 2; see Extended Data Fig. 3 for details), so the temporal window for tunnelling in the single-orbital snapshot image (compare Fig. 2c and Fig. 2e) is determined to be only 115 fs (see Extended Data Fig. 3).

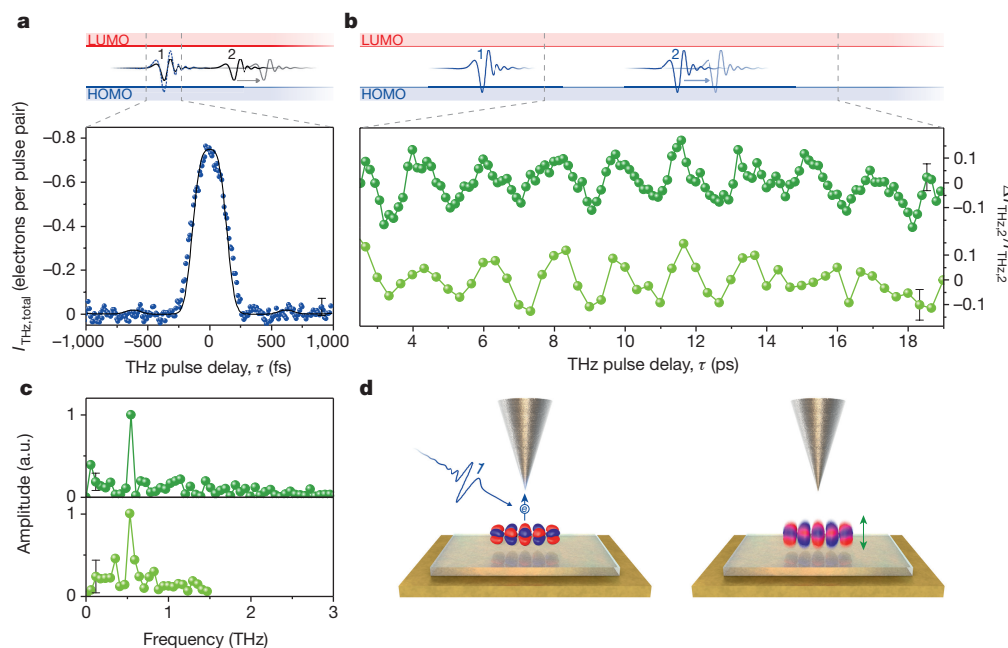
Femtosecond control over the tunnelling time window can be used to capture the ultrafast dynamics of an individual molecule following a single tunnelling event (Fig. 3b). To this end, we introduce a new quantum pump/probe scheme, where two successive, identical THz pulses are individually tuned in resonance with the pentacene HOMO (Fig. 3b, top). Ultrafast dynamics driven by the first field transient (pump) thus potentially modify the rectified current,  $\tilde{I}_{\text{THz},2}$  (see Methods), induced by the second (probe). Remarkably, this probe current measured as a function of the pump–probe delay,  $\tau$  (Fig. 3b), indeed exhibits a marked oscillation at a frequency of 0.5 THz (Fig. 3c).

We interpret this behaviour as a coherent vibration of the molecule, which may be understood as a manifestation of the Franck–Condon principle at the single-molecule level: the pump stimulates tunnelling of an electron from the HOMO to the tip, leaving the molecule temporarily

charged. The abrupt change to the Coulomb and van der Waals forces on the molecule imparts an impulse that prepares a dominantly vertical vibration of the molecular frame over the salt layer<sup>29</sup>. We expect the HOMO to be refilled on the  $\sim 100$ -fs timescale (see Methods), so vibrations occur about the equilibrium position of the neutral molecule. The probe pulse subsequently samples the resulting oscillation of the molecule's position (Fig. 3d), which is well defined because the time window for tunnelling out of the molecule is much shorter than the vibrational period of the relatively heavy molecular frame. Assuming that the  $\pm 10\%$  oscillations of the probe current (Fig. 3b) originate from a modulation of the tunnelling-barrier thickness owing to the molecule's motion, its average oscillation amplitude can be estimated to be  $\pm 4$  picometres from the tunnelling decay rate (current versus barrier thickness).

The oscillatory mode of the molecule is characteristic of the specific van der Waals interaction between pentacene and the substrate. To demonstrate this, we prepared monolayer salt islands at a higher temperature (see Methods), thereby removing the missing-row reconstruction of the Au surface underneath (Fig. 4a, b). Pump/probe measurements recorded over an adsorbed pentacene molecule (at the position indicated by the red circle in Fig. 4c) reveal a molecular oscillation frequency of 0.3 THz (Fig. 4d), which is different from the oscillation frequency observed for pentacene on the previous substrate (Fig. 3b, c). The shift is a direct result of altering the Au reconstruction underneath the decoupling layer, as this modifies the adsorption potential of pentacene. Furthermore, under identical experimental conditions to those in Fig. 4d, we observe a different oscillation frequency (Fig. 4e) for a copper phthalocyanine molecule adsorbed on the same salt island (recorded at the blue circle in Fig. 4c), excluding alternative mechanisms as the origin of the observed current oscillations (see Methods). In each case, we coherently stimulate and resolve



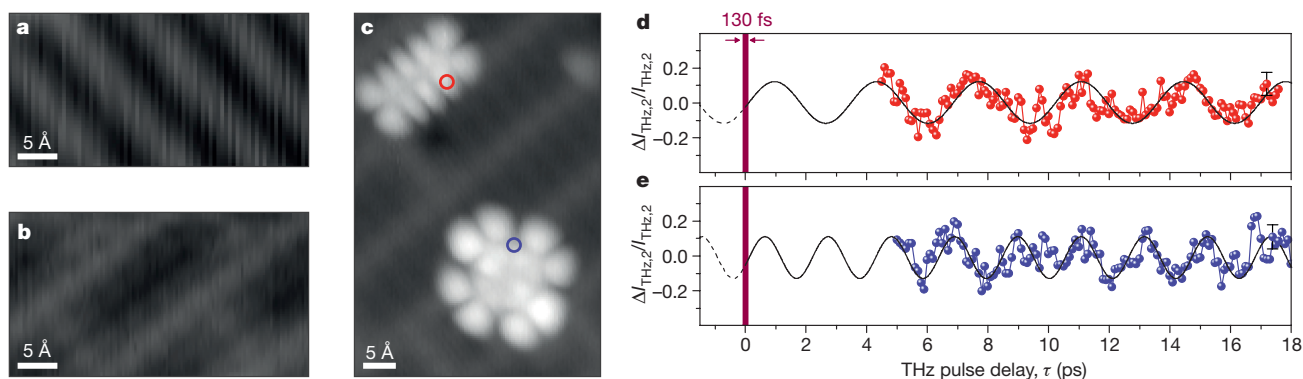


**Figure 3 | Ultrafast dynamics of a single molecule.** **a**, Bottom, THz-pulse autocorrelation for two identical THz pulses, in units of rectified electrons per THz pulse pair, measured with the STM tip located over a pentacene HOMO lobe. Each individual THz pulse (solid line in the diagram above) is too weak (THz far-field amplitude =  $0.12 \text{ kV cm}^{-1}$ ) to enable sequential tunnelling through the HOMO. Upon perfect overlap ( $\tau = 0$ ), however, the resulting transient does reach the HOMO level (dashed line). The experimental  $I_{\text{THz, total}}$  versus  $\tau$  values (dots in graph) are fit (solid line) by a simulation based on rectifying the electro-optically measured THz waveform using a  $dI/dV$  represented by the sum of two Gaussians. The best fit yields a time resolution of 130 fs here and 115 fs for the settings used for imaging. **b**, Measurement of a single pentacene molecule's time dynamics in a pump/probe experiment. Here, each individual THz pulse

is strong enough (THz far-field amplitude =  $0.22 \text{ kV cm}^{-1}$ ) to facilitate HOMO tunnelling. The current due to the second pulse exhibits coherent oscillations as a function of delay time,  $\tau$ . This pump-induced variation,  $\Delta I_{\text{THz}, 2}$ , about the average current,  $I_{\text{THz}, 2}$ , originating from the second pulse is shown (dark green points and curve). A second scan taken on the same molecule confirms a fixed phase relation (light green points and curve). **c**, The Fourier spectra of the measurements in **b** exhibit sharp peaks at  $\sim 0.5 \text{ THz}$ . **d**, Illustration of the pump/probe measurement. The first THz pulse excites a vertical vibration by removing an electron from the HOMO. The second pulse detects the instantaneous height of the oscillating molecule as a change in the net current. Error bars indicate the standard deviation from the fit (**a**), a sinusoidal oscillation (**b**), and the corresponding standard deviation in the frequency domain (**c**).

the oscillatory mode of the molecule in the time domain, and can even trace the oscillations back to identify the absolute phase at the triggering event (see Fig. 4d, e). We find that, for all observed oscillations, the initial phase is identical regardless of the oscillation frequency and is consistent with a 130-fs-long excitation window. In contrast to frequency-based spectroscopy, our coherent pump/probe tunnelling

scheme reveals molecular modes that are normally obscured by incoherent broadening, while simultaneously providing sub-orbital spatial resolution. Molecular motion on the scale of single vibrational quanta can thus be both tracked and manipulated by controlled tunnelling events, where we define the energy, position, and time window of lightwave-driven tunnelling.



**Figure 4 | Influence of substrate surface and molecular species on ultrafast single-molecule dynamics.** **a**, **b**, Constant-current DC-STM images of NaCl/Au(110) islands with (a) and without (b) the missing-row Au(110) reconstruction beneath the monolayer NaCl film. Greyscale range =  $2.0 \text{ \AA}$ . **c**, Constant-current DC-STM image of a pentacene molecule HOMO (top left) and a copper phthalocyanine molecule HOMO (bottom right), which are adsorbed next to one another on the same NaCl monolayer island shown in **b** ( $V_{\text{DC}} = -1.75 \text{ V}$ ,  $I = 0.83 \text{ pA}$ ). Greyscale range =  $3.0 \text{ \AA}$ . The red and blue circles indicate the positions at which the pump/probe dynamics shown in **d**, **e** were measured on

each molecule. **d**, **e**, Quantum pump/probe measurements recorded over the pentacene molecule (**d**, red) and the copper phthalocyanine molecule (**e**, blue). Although the experimental conditions were identical for **d** and **e**, the results show coherent oscillations at different frequencies ( $0.3 \text{ THz}$  and  $0.5 \text{ THz}$ , respectively). Error bars indicate the standard deviation from the sinusoidal fits to the time-domain oscillations (black lines in **d**, **e**), which identify the dominant spectral component in each case. Tracing each fit back to  $\tau = 0$  reveals a consistent absolute phase with a temporal offset of less than 130 fs (vertical purple belt). Sinusoidal fits to the oscillations in Fig. 3b also give the same phase.

The development of ultrafast THz-STM of well-defined quantum states opens the door to a qualitatively new experimental domain in nanoscience. Now, the femtosecond-scale quantum motion of molecules—ranging from vibrations to electronic excitations—can be probed on the scale of a single electronic orbital. Furthermore, the dynamics of the energy landscape associated with single electrons and phonons can be traced directly by ultrafast tunnelling, providing a way of exploring few-particle quantum statistics in single molecules with sub-cycle time resolution. Moreover, combining an optical pump pulse with the THz-STM probe will provide access to a multitude of optical excitations, including those in the time window immediately following the pump pulse ( $\tau < 2$  ps). In the near future, we envisage watching single-electron lightwave electronics in molecular circuits, single-molecule movies, and chemical reactions in four dimensions, visualizing the initial reaction steps of key elementary processes in chemistry and biology.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 10 March; accepted 26 August 2016.**

- Itatani, J. *et al.* Tomographic imaging of molecular orbitals. *Nature* **432**, 867–871 (2004).
- Goulielmakis, E. *et al.* Attosecond control and measurement: lightwave electronics. *Science* **317**, 769–775 (2007).
- Neppl, S. *et al.* Direct observation of electron propagation and dielectric screening on the atomic length scale. *Nature* **517**, 342–346 (2015).
- Hohenleutner, M. *et al.* Real-time observation of interfering crystal electrons in high-harmonic generation. *Nature* **523**, 572–575 (2015).
- Feist, A. *et al.* Quantum coherent optical phase modulation in an ultrafast transmission electron microscope. *Nature* **521**, 200–203 (2015).
- Krüger, M., Schenk, M. & Hommelhoff, P. Attosecond control of electrons emitted from a nanoscale metal tip. *Nature* **475**, 78–81 (2011).
- Wimmer, L. *et al.* Terahertz control of nanotip photoemission. *Nature Phys.* **10**, 432–436 (2014).
- Iwaszczuk, K., Zalkovskij, M., Strikwerda, A. C. & Jepsen, P. U. Nitrogen plasma formation through terahertz-induced ultrafast electron field emission. *Optica* **2**, 116–123 (2015).
- Repp, J., Meyer, G., Stojković, S. M., Gourdon, A. & Joachim, C. Molecules on insulating films: scanning-tunneling microscopy imaging of individual molecular orbitals. *Phys. Rev. Lett.* **94**, 026803 (2005).
- Strosio, J. A. & Celotta, R. J. Controlling the dynamics of a single atom in lateral atom manipulation. *Science* **306**, 242–247 (2004).
- Stipe, B. C., Rezaei, M. A. & Ho, W. Single-molecule vibrational spectroscopy and microscopy. *Science* **280**, 1732–1735 (1998).
- Hamers, R. J. & Cahill, D. G. Ultrafast time resolution in scanned probe microscopies. *Appl. Phys. Lett.* **57**, 2031 (1990).
- Nunes, G., Jr & Freeman, M. R. Picosecond resolution in scanning tunneling microscopy. *Science* **262**, 1029–1032 (1993).
- Kemiktarak, U., Ndikum, T., Schwab, K. C. & Ekinci, K. L. Radio-frequency scanning tunnelling microscopy. *Nature* **450**, 85–88 (2007).
- Terada, Y., Yoshida, S., Takeuchi, O. & Shigekawa, H. Real-space imaging of transient carrier dynamics by nanoscale pump-probe microscopy. *Nature Photon.* **4**, 869–874 (2010).
- Loth, S., Etzkorn, M., Lutz, C., Eigler, D. M. & Heinrich, A. J. Measurement of fast electron spin relaxation times with atomic resolution. *Science* **329**, 1628–1630 (2010).
- Wu, S. W. & Ho, W. Two-photon-induced hot-electron transfer to a single molecule in a scanning tunneling microscope. *Phys. Rev. B* **82**, 085444 (2010).
- Cocker, T. L. *et al.* An ultrafast terahertz scanning tunnelling microscope. *Nature Photon.* **7**, 620–625 (2013).
- Yoshida, S. *et al.* Probing ultrafast spin dynamics with optical pump-probe scanning tunnelling microscopy. *Nature Nanotech.* **9**, 588–593 (2014).
- Merino, P., Große, C., Rosławska, A., Kuhnke, K. & Kern, K. Exciton dynamics of C<sub>60</sub>-based single-photon emitters explored by Hanbury Brown-Twiss scanning tunnelling microscopy. *Nature Commun.* **6**, 8461 (2015).
- Wagner, M. *et al.* Ultrafast and nanoscale plasmonic phenomena in exfoliated graphene revealed by infrared pump-probe nanoscopy. *Nano Lett.* **14**, 894–900 (2014).
- Eisele, M. *et al.* Ultrafast multi-terahertz nano-spectroscopy with sub-cycle temporal resolution. *Nature Photon.* **8**, 841–845 (2014).
- Gao, M. *et al.* Mapping molecular motions leading to charge delocalization with ultrabright electrons. *Nature* **496**, 343–346 (2013).
- Zewail, A. H. Four-dimensional electron microscopy. *Science* **328**, 187–193 (2010).
- Jahng, J. *et al.* Ultrafast pump-probe force microscopy with nanoscale resolution. *Appl. Phys. Lett.* **106**, 083113 (2015).
- Zhang, R. *et al.* Chemical mapping of a single molecule by plasmon-enhanced Raman scattering. *Nature* **498**, 82–86 (2013).
- Lee, J., Perdue, S. M., Perez, A. R. & Apkarian, V. A. Vibronic motion with joint angstrom–femtosecond resolution observed through Fano progressions recorded within one molecule. *ACS Nano* **8**, 54–63 (2014).
- Zhang, Y. *et al.* Visualizing coherent intermolecular dipole–dipole coupling in real space. *Nature* **531**, 623–627 (2016).
- Park, H. *et al.* Nanomechanical oscillations in a single-C<sub>60</sub> transistor. *Nature* **407**, 57–60 (2000).
- Yoshida, K., Shibata, K. & Hirakawa, K. Terahertz field enhancement and photon-assisted tunneling in single-molecule transistors. *Phys. Rev. Lett.* **115**, 138302 (2015).

**Acknowledgements** We thank A. Pöllmann, F. Albrecht, R. Kronfeldner, M. Eisele and M. Furthmeier for assistance, and F. Evers, F. A. Hegmann, V. Jelic, P. C. M. Planken, M. Grifoni and K. Richter for discussions. We acknowledge financial support from the Volkswagen Foundation (Lichtenberg program), the European Research Council through grant number 305003 (QUANTUMsubCYCLE), and the Deutsche Forschungsgemeinschaft (DFG) through GRK 1570 and research grants HU1598/3 and CO1492/1. T.L.C. thanks the A. v. Humboldt Foundation.

**Author Contributions** T.L.C., D.P., P.Y., J.R. and R.H. conceived, set up and carried out the experiments. T.L.C., D.P., J.R. and R.H. analysed the data and wrote the manuscript.

**Additional Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.R. ([jascha.repp@physik.uni-regensburg.de](mailto:jascha.repp@physik.uni-regensburg.de)) or R.H. ([rupert.huber@physik.uni-regensburg.de](mailto:rupert.huber@physik.uni-regensburg.de)).

**Reviewer Information** *Nature* thanks H. Shigekawa and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

**THz optical set-up.** Intense, phase-locked THz pulses are generated by tilted-pulse-front optical rectification of femtosecond near-infrared pulses (centre wavelength 1,028 nm, pulse duration (full width at half-maximum) 250 fs) from a regenerative ytterbium-doped potassium gadolinium tungstate (Yb:KGW) laser amplifier (repetition rate 0.61 MHz) in lithium niobate. The waveform of the generated single-cycle THz transients is directly measured by electro-optic sampling in a 0.5-mm-thick (110)-cut zinc telluride crystal using femtosecond near-infrared gate pulses (centre wavelength 800 nm, pulse duration 10 fs). Pairs of mutually delayed THz transients are prepared by transmitting the THz pulses through a Michelson interferometer, in which the computer-controlled position of one end mirror sets the delay time,  $\tau$ . The field amplitude can be continuously scaled without changing the waveform using wire-grid polarizers. We note that the precise field enhancement at the tip depends on the local tip preparation for a specific experiment, but is constant for a given tip position as long as the tip shape is maintained.

**STM set-up and sample preparation.** The homebuilt STM is based on a Besocke design<sup>31</sup> and operates in ultrahigh vacuum (UHV; pressure  $\sim 7 \times 10^{-11}$  mbar) and at low temperatures down to 11 K. The bias is applied to the sample, and the bias voltage window refers to the energy region between the tip and sample Fermi levels. A homebuilt, low-temperature ( $\sim 100$  K), high-gain ( $G = 2.5 \times 10^{10}$ ) preamplifier is used that is mounted in close proximity to the STM head. THz radiation enters the vacuum chamber through a sapphire viewport as a collimated beam and is focused onto the tip by a parabolic mirror that is fixed to the STM scan unit. NaCl is evaporated thermally onto the clean Au(110) sample under UHV conditions at a sample temperature of  $\sim 295$  K (missing-row reconstruction) and  $\sim 400$  K (no missing-row reconstruction), in two different preparations. Pentacene or copper phthalocyanine molecules are deposited onto the cold sample ( $T < 15$  K) while it is located inside the STM.

**THz-STM measurements.** To detect the THz-induced current, we modulate the THz pulses at 475 Hz with an optical chopper (faster than the STM feedback loop bandwidth), and measure the AC tunnel current with lock-in detection. We note that the average current due to the THz pulse train is small relative to the DC setpoint current, which is typically of the order of 1 pA, and can be in the opposite direction. The average THz-induced current,  $I_{\text{THz}}$ , in units of amps, can be converted into the number of electrons per pulse by:  $I_{\text{THz}} [e^- \text{ pulse}^{-1}] = I_{\text{THz}} [\text{A}] / (ef)$ , where  $f$  is the repetition rate of the THz pulses and  $e$  is the elementary charge. For example, if only one THz pulse is produced per laser shot, 98 fA corresponds to 1 electron per pulse. The combination of our low-noise preamplifier, laser repetition rate, and lock-in detection therefore allows us to measure THz-induced currents down to 0.01 electrons per pulse.

In an autocorrelation measurement (for example, Fig. 3a), each generated THz pulse is split in two and separated by a variable delay time,  $\tau$ , in the Michelson interferometer. The peak fields of the two identical THz pulses focused onto the tip are adjusted such that the combined peak field reaches the HOMO only when the pulses fully overlap ( $\tau = 0$ ). The optical generation beam is modulated, so the measured AC current corresponds to the total THz-induced current created by the two pulses.

Conversely, pump/probe measurements (for example, Figs. 3b and 4d, e) are made by modulating one of the two THz pulse trains in a pair independently (that is, one arm of the Michelson interferometer contains a chopper). Additionally, in quantum pump/probe measurements, we adjust the peak field of each THz pulse in a pair such that the pulses access a particular orbital independently when separated (that is, for large values of  $\tau$ ). We then compare the AC tunnel current,  $I_{\text{THz},1}$ , induced by the first pulse only, the corresponding current  $I_{\text{THz},2}$  caused solely by the second pulse, and the current  $I_{\text{THz},\text{total}}(\tau)$ , measured when both THz fields act on the junction. Finally, we define  $I_{\text{THz},2}(\tau) = I_{\text{THz},\text{total}}(\tau) - I_{\text{THz},1}$  and  $\Delta I_{\text{THz},2}(\tau) = \tilde{I}_{\text{THz},2}(\tau) - I_{\text{THz},2}$ .

In the measurements shown here, we avoid the range  $\tau < 2$  ps, because of interference. To obtain a detectable THz-induced current, the tunnelling rate between tip and molecule has to be large enough to ensure that an electron has an appreciable probability of tunnelling during the ultrashort tunnelling time window at the THz peak. This high tunnelling rate is facilitated by a close tip–molecule distance and will—outside the THz peak—lead to an electron tunnelling from

the tip, thereby refilling the HOMO at a similar rate. This process competes with the HOMO being refilled from the substrate, where extrapolation of the results in ref. 32 indicates a tunnelling time of the order of 100 fs through a NaCl monolayer. The net tunnel current that is observed stems from those instances only for which the HOMO is refilled from the substrate. Signals similar to the oscillations shown for  $\tau > 2$  ps are observed at negative delay times, when the roles of the identical THz pump and probe pulses are interchanged. Finally, we note that we observe a small tip expansion owing to heating by the THz pulse train, but it is on the order of 10 picometres and occurs on a timescale of seconds (which is far slower than our chopping frequency) and therefore has no effect on our measurements.

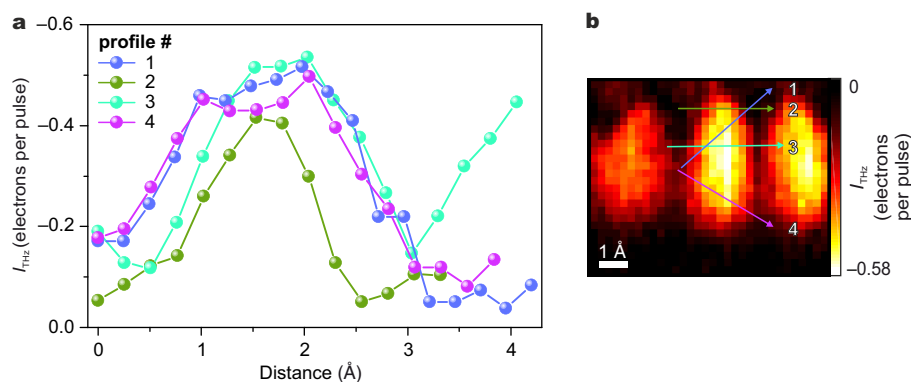
**Simulations.** To simulate the spatial distribution of the HOMO in the THz-STM images (Fig. 2d, i), we adopt the Bardeen model<sup>33</sup>, where the tunnel current is proportional to the square of the matrix element between the relevant tip and sample wavefunctions. As in the Tersoff–Hamann approach<sup>33</sup>, we approximate the tip wavefunction as an s-wave, whereas we use the DFT-derived HOMO of the free molecule as the sample wavefunction (calculated using the ‘TURBOMOLE’ computational framework<sup>34</sup> in version 6.2, with the B3-LYP functional in the def2-TZVP basis set). To model the spatial dependence of the tunnel current, we evaluate the convolution of the HOMO with the s-wave tip for the experimentally measured topography (Fig. 2b). The best agreement between the simulations and the THz-STM images is obtained for s-waves with 1/e decay lengths of  $\leq 0.5$  Å, which is consistent with the spatial resolution of the experimental image (0.6 Å; see Extended Data Fig. 1).

**Origin of different oscillation frequencies.** For pentacene molecules adsorbed on NaCl monolayer islands on Au(110) with the missing-row reconstruction, we observed a consistent oscillation frequency of 0.5 THz (Fig. 3). We also investigated copper phthalocyanine (CuPc) molecules on the same substrate and observed a higher oscillation frequency of 0.7 THz (data not shown). For pentacene and CuPc adsorbed on NaCl islands on Au(110) without the missing-row reconstruction, we found oscillation frequencies of 0.3 THz (Fig. 4d) and 0.5 THz (Fig. 4e), respectively. The higher frequency observed for CuPc molecules is consistent with the slightly polar character of the bonds inside CuPc, which results in an increased electrostatic interaction with the underlying ionic NaCl film that stiffens the binding to the substrate. Generally, though, the weak interaction of a relatively large  $\pi$ -conjugated organic molecule with an insulating substrate is very difficult to model and to predict with theory or corresponding simulations.

On the experimental side, vibrational frequencies of 1.5–2 THz have been reported for hydrocarbons directly adsorbed on Cu(100)<sup>35</sup> and Ru(0001)<sup>36</sup>. It is expected that the interaction of the  $\pi$ -conjugated molecules with the substrate is substantially weakened by introducing an insulating film. Hence, for pentacene and CuPc on NaCl/Au(110), frequencies well below 1.5 THz are expected, in qualitative agreement with our observations. Most importantly, the occurrence of different oscillator frequencies for different molecular species and different surface reconstructions with otherwise identical conditions rules out many other potential causes of the observed oscillations: namely, plasmonic resonances of the junction; THz resonances of the entire tip that influence the THz coupling; interference effects of potential trailing oscillations of the THz waveform; and coherently excited surface modes that change the substrate–molecule tunnelling probability as a function of  $\tau$ .

1. Meyer, G. A simple low-temperature ultrahigh-vacuum scanning tunneling microscope capable of atomic manipulation. *Rev. Sci. Instrum.* **67**, 2960–2965 (1996).
2. Steurer, W., Gross, L. & Meyer, G. Local thickness determination of thin insulator films via localized states. *Appl. Phys. Lett.* **104**, 231606 (2014).
3. Chen, C. J. *Introduction to Scanning Tunneling Microscopy* (Oxford Univ. Press, 1993).
4. Ahlrichs, R., Bär, M., Häser, M., Horn, M. & Kölmel, C. Electronic structure calculations on workstation computers: the program system TURBOMOLE. *Chem. Phys. Lett.* **162**, 165–169 (1989).
5. Witte, G. & Wöll, C. External vibrations of hydrocarbons on Cu(100). *J. Chem. Phys.* **103**, 5860–5863 (1995).
6. Witte, G. et al. Damping of molecular motion on a solid substrate: evidence for electron-hole pair creation. *Phys. Rev. Lett.* **80**, 121–124 (1998).

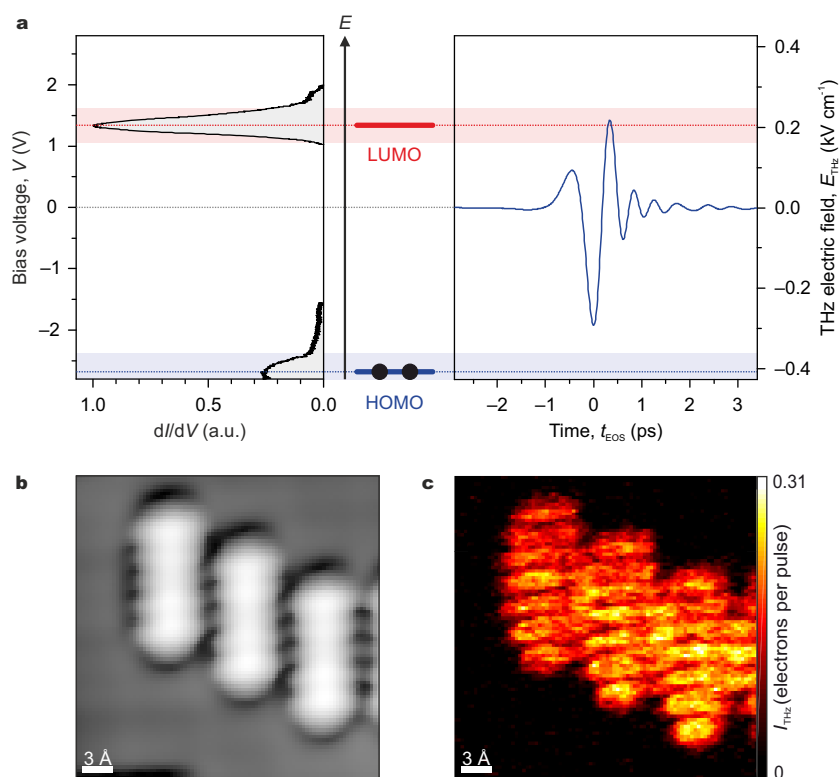




### Extended Data Figure 1 | Spatial resolution of THz-STM imaging.

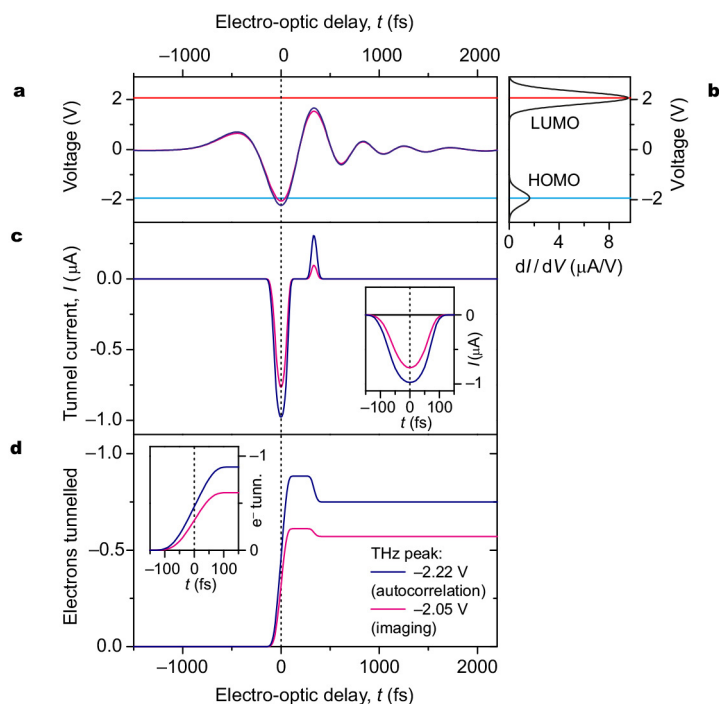
**a, b**, Line-scan cuts (**a**) through a constant-height, zero-bias THz-STM image (**b**), where the THz-induced current  $I_{\text{THz}}$  is shown in units of rectified electrons per THz pulse. The pixel size in the image is  $0.25 \text{ \AA} \times 0.25 \text{ \AA}$  and the white scale bar is  $1 \text{ \AA}$  long. The image contains edges where the signal rises from the background level to a maximum within  $0.75 \text{ \AA}$ . The signal rise from 10% to 90% in these cases occurs over approximately  $0.6 \text{ \AA}$ , which provides an upper bound for the spatial resolution of the image. This estimate agrees with simulations of the

spatial distribution of the rectified current (Fig. 2d, i). There, we calculate the square of the matrix element between the pentacene HOMO and the tip wavefunction (Bardeen model). The tip wavefunction, which ultimately determines the attainable spatial resolution, is modelled as an s-wave (as in the Tersoff–Hamann approach) and the spatial decay of this wavefunction is used as a fitting parameter, as is the tip height. The best agreement between the experimental and simulated images is found for s-waves with decay lengths of  $\leq 0.5 \text{ \AA}$ , consistent with our spatial resolution estimate of  $0.6 \text{ \AA}$  based on the line scans in **a**.



**Extended Data Figure 2 | Ultrafast THz-STM imaging of the lowest unoccupied orbital of pentacene.** **a**, When Au(110) is replaced with Cu(100) as the substrate, the HOMO and LUMO levels realign with respect to the Fermi level of the substrate in such a way that the LUMO can be probed by THz-induced currents. This situation is depicted schematically here, in analogy to Fig. 1b. Left, steady-state differential conductance measured on a pentacene molecule adsorbed on NaCl/Cu(100). Right, THz electric-field waveform measured by electro-optic sampling. Although the THz waveform is the same as that in Fig. 1b, it now induces sequential tunnelling into the pentacene LUMO at the most intense positive half-cycle, while the negative half-cycle is too weak to allow for HOMO tunnelling. **b**, Steady-state constant-current STM image

of three pentacene molecules side-by-side at  $V_{\text{DC}} = 100$  mV and  $I = 1.4$  pA. Greyscale range =  $2.4 \text{ \AA}$ . **c**, Constant-height THz-STM image of the same sample area as in **b**, with  $V_{\text{DC}} = 0$  mV and the maximum positive THz voltage set to  $V_{\text{THz}}^{\text{peak}} \approx +1.3$  V. The spatial distribution of the THz-induced current closely resembles the LUMO density for each molecule, indicating state-selective THz-induced LUMO tunnelling. The onset field for this process agrees with the modified alignment of the orbitals. Specifically, the peak THz voltage in the positive direction that is predicted based on the field enhancement determined from HOMO tunnelling is consistent with the new LUMO transport level voltage observed in the steady-state  $dI/dV$  curve here. This agreement lends further support to the proposed mechanism for THz-STM in this regime.



**Extended Data Figure 3 | Modelling ultrafast terahertz-induced tunnelling out of a pentacene HOMO. a,** Terahertz voltage waveforms used

**tunnelling out of a pentacene HOMO.** **a**, Terahertz voltage waveforms used in the simulations. We approximate the time trace of the THz voltage with the THz electric-field waveform measured in the far field by electro-optic sampling. In all plots, the pink curve corresponds to a waveform with a peak of  $-2.05$  V and the navy-blue curve corresponds to a waveform with a peak of  $-2.22$  V. The scaling factor and resulting peak voltages are determined by fitting the shape of the experimental autocorrelation (Fig. 3a) as follows. We simulate the autocorrelation by taking the sum of two THz pulses and calculating the total rectified current from the resulting transient as a function of the delay time between the pulses. **b**, The  $dI/dV$  characteristic needed for the simulation is obtained by modelling the HOMO and LUMO molecular-resonance peaks in the experimental  $dI/dV$  curve (Fig. 2b, centre) with two Gaussians. This  $dI/dV$  curve is then scaled up to account for the significantly lower tip height for THz-STM imaging. Note that the scaling factor does not affect the shape of the autocorrelation. Blue line, centre of HOMO Gaussian,  $-1.93$  V ( $0.62$  V full width at half-maximum, FWHM). Red line, centre of LUMO Gaussian,  $2.06$  V ( $0.58$  V FWHM). **c**, Resulting current response induced by the THz voltage waveform in **a** when applied to a junction with a  $dI/dV$  relation defined by **b**. The asymmetry of the THz voltage pulse leads to a much larger current response in the negative bias direction than in the positive bias direction. Inset, induced current response during the negative crest of the THz voltage waveform. Pink curve,  $120$ -fs FWHM; navy-blue curve,  $140$ -fs FWHM. **d**, Total number of electrons that have tunneled across the junction, calculated from the integral of **c**. Negative numbers refer to the negative bias direction, that is, electron tunnelling from the HOMO to the tip. Inset, rise of the rectified electron signal during the negative crest of the THz voltage waveform. The rise time from 10% to 90% of the maximum signal is  $115$  fs for the pink curve and  $130$  fs for the navy-blue curve. The simulated autocorrelation for the  $-2.22$  V peak (navy-blue curve) provides the best fit to the shape of the measured autocorrelation. The  $dI/dV$  curve is scaled such that the autocorrelation peak at  $-2.22$  V matches the measured peak of approximately  $-0.75$  electrons per THz pulse. The peak voltage of the THz pulses used in the imaging configuration is then determined by finding the THz peak for which the simulations yield this number of rectified electrons per THz pulse. For example, in Fig. 2e the maximum observed signal is approximately  $-0.58$  electrons per THz pulse, and the best agreement is found for  $-2.05$  V (pink curve). We note that the simulations are based on the assumption that the THz pulse modulates the bias of the junction quasi-instantaneously. In our simulations we disregard any blocking of tunnelling or other effects resulting from the finite time until an electron from the substrate refills the molecular state. This blocking is expected to become important if the THz current reaches or even exceeds one electron per pulse.



# Catalytic alkylation of remote C–H bonds enabled by proton–coupled electron transfer

Gilbert J. Choi<sup>1</sup>, Qilei Zhu<sup>1\*</sup>, David C. Miller<sup>1\*</sup>, Carol J. Gu<sup>1</sup> & Robert R. Knowles<sup>1</sup>

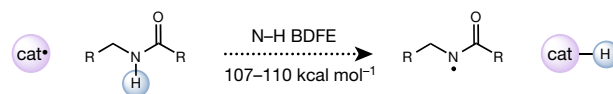
Despite advances in hydrogen atom transfer (HAT) catalysis<sup>1–5</sup>, there are currently no molecular HAT catalysts that are capable of homolysing the strong nitrogen–hydrogen (N–H) bonds of *N*-alkyl amides. The motivation to develop amide homolysis protocols stems from the utility of the resultant amidyl radicals, which are involved in various synthetically useful transformations, including olefin amination<sup>6–11</sup> and directed carbon–hydrogen (C–H) bond functionalization<sup>12–16</sup>. In the latter process—a subset of the classical Hofmann–Löffler–Freitag reaction—amidyl radicals remove hydrogen atoms from unactivated aliphatic C–H bonds<sup>17–21</sup>. Although powerful, these transformations typically require oxidative *N*-prefunctionalization of the amide starting materials to achieve efficient amidyl generation. Moreover, because these *N*-activating groups are often incorporated into the final products, these methods are generally not amenable to the direct construction of carbon–carbon (C–C) bonds. Here we report an approach that overcomes these limitations by homolysing the N–H bonds of *N*-alkyl amides via proton–coupled electron transfer. In this protocol, an excited-state iridium photocatalyst and a weak phosphate base cooperatively serve to remove both a proton and an electron from an amide substrate in a concerted elementary step. The resultant amidyl radical intermediates are shown to promote subsequent C–H abstraction and radical alkylation steps. This C–H alkylation represents a catalytic variant of the Hofmann–Löffler–Freitag reaction, using simple, unfunctionalized amides to direct the formation of new C–C bonds. Given the prevalence of amides in pharmaceuticals and natural products, we anticipate that this method will simplify the synthesis and structural elaboration of amine-containing targets. Moreover, this study demonstrates that concerted proton–coupled electron transfer can enable homolytic activation of common organic functional groups that are energetically inaccessible using traditional HAT-based approaches.

Owing to their exceptional homolytic stability, there are currently no reported HAT catalysts that are capable of activating the strong N–H bonds of 2° *N*-alkyl amides (Fig. 1a). Here we report a protocol that addresses this deficit, selectively homolysing these bonds through a concerted proton–coupled electron transfer (PCET) event. In these reactions, an excited-state iridium oxidant and a weak phosphate base cooperatively serve to remove both a proton and an electron from an amide substrate in a single elementary step to furnish a reactive amidyl radical. This method of amide activation was then exploited to develop a catalytic variant of the classical Hofmann–Löffler–Freitag reaction (Fig. 1b) that uses simple amide substrates to direct the formation of new C–C bonds (Fig. 1c) with unfunctionalized alkanes. The design, optimization and mechanistic study of this process are reported herein.

Our initial efforts sought to identify conditions that are effective for the direct C–H alkylation of model amide **1**, which has a N–H bond-dissociation free energy (BDFE) of 107 kcal mol<sup>–1</sup> (refs 22, 23).

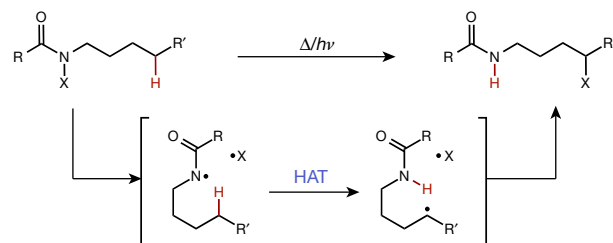
In analogy to our previous work on *N*-aryl amide activation<sup>6,7</sup>, we envisioned a catalytic cycle wherein the phosphate base would first associate via hydrogen bonding to the amide N–H bond of **1** (Fig. 2). Oxidation of this hydrogen-bond complex by the excited state of an iridium photocatalyst would result in formal homolysis of the strong N–H bond

## a Challenges in catalytic homolysis of strong N–H bonds



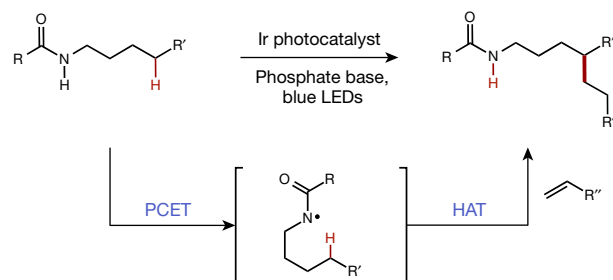
No known catalysts for selective homolysis of *N*-alkyl amide N–H bonds

## b Classical Hofmann–Löffler–Freitag reactions



• *N*-functionalization required • No methods for C–C bond formation

## c Catalytic C–H alkylation enabled by PCET

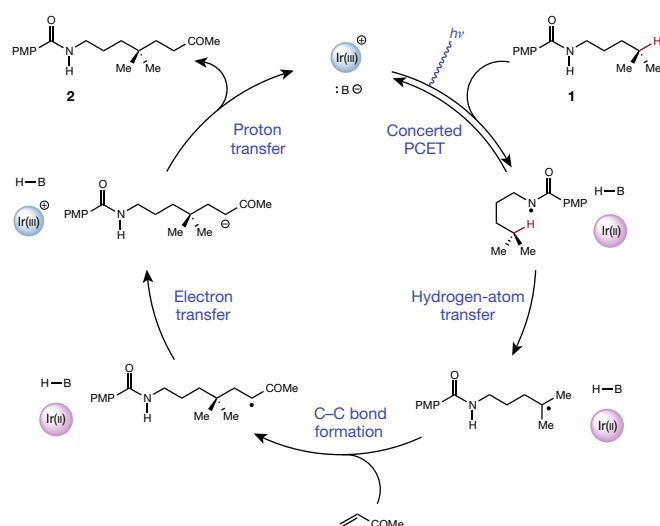


Directed alkylation of remote C–H bonds enabled by amide PCET

**Figure 1 | Design and development of a catalytic amidyl-mediated C–H alkylation.** **a**, With bond-dissociation free energies (BDFEs) of 107–110 kcal mol<sup>–1</sup>, there are no reported molecular catalysts capable of homolysing the N–H bonds of *N*-alkyl amides. **b**, The classical Hofmann–Löffler–Freitag reaction enables the selective abstraction of C–H bonds at positions remote from the amidyl radical via hydrogen-atom transfer (HAT).  $\Delta$ , heat;  $h\nu$ , light. **c**, Proposed direct C–H alkylation of remote C–H bonds via an intermediate amidyl radical generated by concerted oxidative proton–coupled electron transfer (PCET).

<sup>1</sup>Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA.

\*These authors contributed equally to this work.

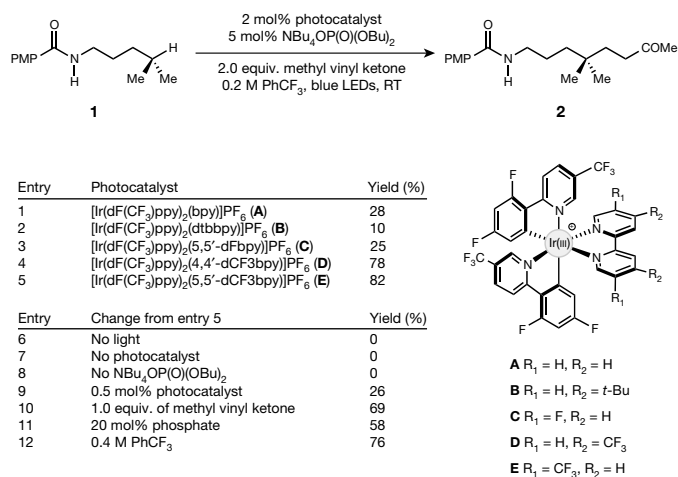


**Figure 2 | Proposed catalytic cycle.** The catalytic cycle begins with an association of the phosphate base (B) via hydrogen bonding to the amide N–H bond of the substrate **1**. Oxidative proton-coupled electron transfer (PCET) generates a neutral amidyl radical. A 1,5-hydrogen-atom abstraction then occurs to generate the distal carbon-centred radical. This intermediate undergoes a conjugate addition with an olefin acceptor to furnish a new C–C bond and an  $\alpha$ -carbonyl radical. Electron transfer from the reduced Ir(II) catalyst furnishes an enolate. Proton transfer from phosphoric acid produces the product **2** and returns the catalysts to their active forms.

via concerted PCET<sup>24–26</sup> and furnish a reactive amidyl radical. This oxidation would then be relayed via 1,5-hydrogen-atom abstraction from a distal aliphatic C–H bond through a cyclic transition state, resulting in site-selective formation of a new alkyl radical. In turn, this open-shell intermediate could engage in a conjugate addition reaction with an electron-deficient olefin partner to furnish a new C–C bond and an  $\alpha$ -carbonyl radical. Electron transfer to this electrophilic radical from the reduced Ir(II) form of the photocatalyst would generate an enolate anion that could be promptly protonated by the phosphoric acid to provide the C–H alkylation product **2** and return both catalysts to their active forms.

Using the catalyst system that proved optimal in our previous studies for *N*-aryl amidyl generation—[Ir(dF(CF<sub>3</sub>)ppy)<sub>2</sub>(bpy)]PF<sub>6</sub> (ref. 27) and the tetrabutyl ammonium salt of dibutyl phosphate—we found that the desired C–H alkylation product **2** was observed in 28% yield when the reaction was carried out in trifluorotoluene (Fig. 3, entry 1). Seeking to improve on this result, we evaluated several related iridium photocatalysts (entries 2–5) and observed that several gave more promising results. The largest increase in reaction efficiency was observed when [Ir(dF(CF<sub>3</sub>)ppy)<sub>2</sub>(5,5′-dCF<sub>3</sub>bpy)]PF<sub>6</sub> (**E**) was used (excited-state reduction potential  $E_{1/2}^* = 1.30$  V versus Fc/Fc<sup>+</sup> in MeCN), furnishing the desired C–H alkylation product in 82% yield. Control experiments run in the absence of light, photocatalyst and base all provided no detectable product (entries 6–8). Both lower photocatalyst and olefin acceptor loadings still delivered serviceable yields of product (entries 9 and 10), although the yields decreased when higher concentrations of phosphate were used (entry 11). Finally, the model reaction also proceeds well when the reaction was run at higher concentrations (0.4 M) in PhCF<sub>3</sub> (entry 12).

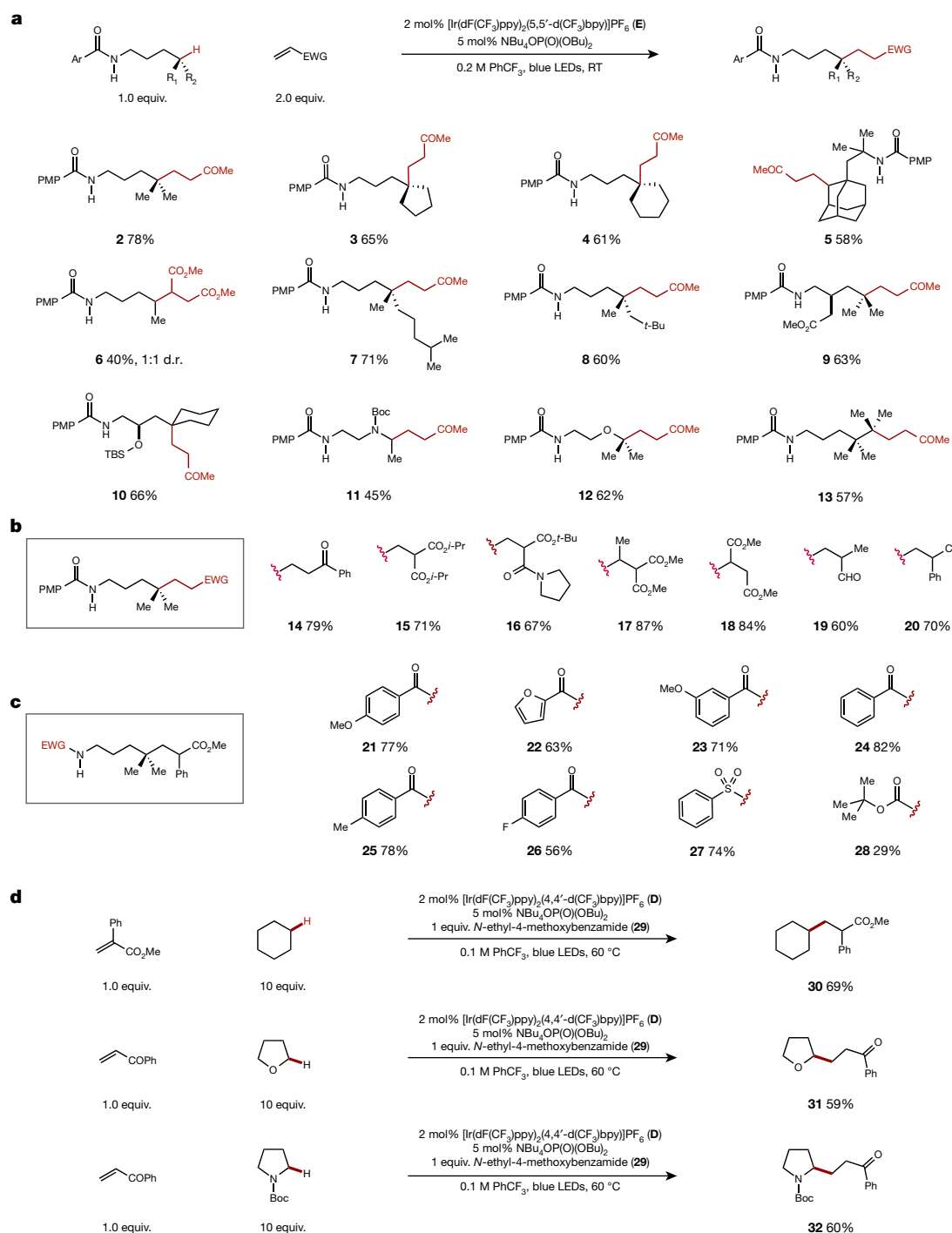
With these optimized conditions, we next evaluated the scope of the C–H alkylation protocol (Fig. 4a–c). On a preparative scale, model amide **1** delivered alkylated adduct **2** in 78% isolated yield. Related substrates, wherein the abstracted hydrogen originates from carbocycles of various sizes, were also alkylated efficiently to form products **3** and **4**. This method is not limited to the abstraction of tertiary methine C–H bonds; secondary methylene-containing substrates were also viable



**Figure 3 | Optimization studies.** The reaction optimization in trifluorotoluene (PhCF<sub>3</sub>) shows that [Ir(dF(CF<sub>3</sub>)ppy)<sub>2</sub>(5,5′-dCF<sub>3</sub>bpy)]PF<sub>6</sub> offers the highest yield of the desired product. NBu<sub>4</sub>P(O)(O)Bu<sub>2</sub> is the tetrabutyl ammonium salt of dibutyl phosphate. Control experiments show that both catalysts are required for product formation along with the use of visible light. Variations from the standard reaction conditions show decreased reaction efficiencies. bpy, bipyridine; dtbbpy, 4,4′-bis(*tert*-butyl) bipyridine; dF(CF<sub>3</sub>)ppy, 3,5-difluoro-2-[5-(trifluoromethyl)-2-pyridinyl] phenyl; 5,5′-dCF<sub>3</sub>bpy, 5,5′-bis(trifluoromethyl) bipyridine.

abstraction partners, as demonstrated by the generation of adamantyl derivative **5**. This example also demonstrates that steric bulk adjacent to the amidyl nitrogen can be tolerated in these reactions. In addition, alkylation was possible at secondary methylene sites on linear alkane substrates (albeit with lower efficiency), as demonstrated by the formation of adduct **6**. The fumarate acceptor in this example was chosen to inductively discourage over-alkylation of the product. The reaction was also successful with amides derived from common terpene feedstocks, generating alkylated product **7**. Although this substrate contains two electronically similar methine C–H bonds, the alkylation event occurs exclusively at the site proximal to the amide. This selectivity would probably be diminished or reversed using conventional intermolecular hydrogen-atom abstractor catalysts, highlighting the utility of this directed approach that is afforded by amide PCET. Also, it was found that steric bulk adjacent to the abstracted site could be tolerated, such as the neopentyl group of product **8**. More complex products such as pregabalin derivative **9** and protected amino alcohol **10** can also be accessed using the reported protocol. Similarly,  $\alpha$ -amino and  $\alpha$ -oxy C–H bonds were operable sites for abstraction, as shown by the formation of alkylated carbamate **11** and ether **12**. In **11**, the methylenes flanking either side of the *N*-Boc substituent and the methylene adjacent to the amide N–H bond are electronically similar, which would limit the selectivity of HAT reactions carried out using conventional intermolecular abstractors. By contrast, the cyclic transition states required for Hofmann–Löffler–Freitag reactivity ensure that only a single alkylated product is observed. Lastly, we observed that efficient 1,6 abstraction occurs when the 1,5 pathway is precluded, as shown in the formation of **13** in 57% yield using catalyst **D**.

With respect to the olefinic partner in these reactions, both alkyl and aryl enone derivatives could be used effectively to furnish adducts such as **2** and **14**. However, reactions with simple acrylate or acrylamide coupling partners provided only traces of C–H alkylation products. We reasoned that this outcome was probably due to the more difficult reduction of the  $\alpha$ -carbonyl radical in these systems (reduction potential  $E_{1/2}^* = -1.37$  V versus Fc/Fc<sup>+</sup>)<sup>28</sup>, which is endergonic relative to the reducing capacity of the Ir(II) state of catalyst **E** ( $E_{1/2}^* = -1.07$  V versus Fc/Fc<sup>+</sup>). To overcome this limitation, we found that various dicarbonyl compounds could be used successfully (**15–17**), including those bearing alkyl substituents at the  $\beta$ -position. These compounds are more



**Figure 4 | Substrate scope.** **a**, A range of amide substrates (**2–13**) can be alkylated in an intramolecular fashion. **b**, The scope of the olefin acceptor illustrates that various common functional groups (**14–20**) can be incorporated into the structure of the model compound (boxed). **c**, Several structurally and electronically distinct electron-withdrawing groups (EWGs; **21–28**) can be incorporated into the structure of the model compound (boxed), including aryl sulfonamides and *N*-Boc carbamates. Note that photocatalyst **D** (instead of **E**) was used in reactions to form products **13** and **27**. **d**, Intermolecular C–H alkylations can also be effected with an excess of the alkane relative to the olefin acceptor. RT, room temperature; Boc = *tert*-butyl carbamate; Me, methyl; *t*-Bu, *tert*-butyl; *i*-Pr, isopropyl; Ph, phenyl; PMP, *p*-methoxyphenyl; TBS, *tert*-butyldimethylsilyl; dF(CF<sub>3</sub>)ppy, 3,5-difluoro-2-[5-(trifluoromethyl)-2-pyridinyl]phenyl; 5,5'-bis(trifluoromethyl)bi-pyridine.

reactive acceptors for the alkyl radical and are amenable to subsequent decarboxylation to furnish formal acrylate and acrylamide addition products. We also found that fumarate derivatives could be alkylated efficiently to furnish 1,2-disubstituted dicarboxylic acid derivative **18**. Various olefins bearing  $\alpha$ -substituents were alkylated efficiently, including methacrolein,  $\alpha$ -phenyl acrylonitrile and  $\alpha$ -phenyl methacrylate, to furnish products **19–21**. Several structurally and electronically distinct benzamide derivatives were alkylated successfully using  $\alpha$ -phenyl acrylate as the olefin acceptor (**21–26**). Additionally, we found that aryl sulfonamides served as hydrogen-atom abstractors using catalyst **D** under otherwise standard conditions, generating **27** in 74% yield. Common *N*-Boc carbamates were also alkylated, albeit with lower efficiency (**28**). This example demonstrates that a conjugated aryl group is not a requirement for amide PCET reactivity.

We found that this method could also be adapted for use in intermolecular C–H functionalizations (Fig. 4d). Specifically, the comparatively strong C–H bonds in cyclohexane (10 equiv.) could be alkylated with  $\alpha$ -phenyl methacrylate in 69% isolated yield using *N*-ethyl-4-methoxybenzamide (**29**) as the abstractor and 2 mol% of iridium catalyst **D** at 60 °C (**30**). The alkylation of tetrahydrofuran and *N*-Boc pyrrolidine also proceeded with good efficiency (**31**, **32**). These results suggest that *N*-alkyl amides have the potential to serve as highly reactive and structurally modular catalysts for radical C–H functionalization.

To evaluate the possible role of PCET in N–H activation, we examined the mechanism of amidyl formation using Stern–Volmer assays and *N*-ethyl-4-methoxybenzamide (**29**) as a model substrate. Dichloromethane was used as a solvent to ensure full solubility of all



reaction components. This is an effective reaction solvent, delivering **21** (Fig. 4c) in 80% yield under otherwise standard conditions. These luminescence quenching experiments revealed that **29** (half-peak potential  $E_p = +1.48$  V versus  $\text{Fc}/\text{Fc}^+$  in MeCN) does not quench the excited state of photocatalyst **E** ( $E_{1/2} = +1.30$  V versus  $\text{Fc}/\text{Fc}^+$  in MeCN). However, solutions containing varying concentrations of **29** and a constant amount of tetrabutyl ammonium dibutyl phosphate resulted in a large decrease in the emission intensity. The quenching was found to be linear with respect to the amide concentration (Stern–Volmer constant  $k_{SV} = 46 \text{ M}^{-1}$ ), consistent with a first-order kinetic dependence. These results rule out stepwise activation of the N–H bond by an electron-transfer–proton-transfer mechanism, because catalyst quenching cannot be effected by the substrate alone.

Next, we observed that the phosphate base alone also quenched the excited state of iridium, although the concentration dependence was nonlinear and saturated at higher concentrations. Further experiments wherein the phosphate concentration was varied in the presence of a constant amount of amide exhibited similar saturation behaviour, but with greater overall quenching efficiency. Seeking to understand these results, we found that the  $^1\text{H}$  NMR chemical shift of the  $\text{C}_3$  proton of the  $\text{dCF}_3\text{bpy}$  ligand is sensitive to the concentration of added phosphate, undergoing a downfield chemical shift of 1.4 p.p.m. in the presence of equimolar phosphate. This observation was attributed to hydrogen bonding between the phosphate and the polarized C–H bonds of the  $\text{dCF}_3\text{bpy}$  ligand of catalyst **E**. Precedent for such an effect is provided by observations<sup>29</sup> of similar effects for ruthenium polypyridyl complexes in dichloromethane in the presence of exogenous anions. We further observed that the presence of 10 equiv. of phosphate does not substantially alter the  $\text{Ir(II/III)}$  redox couple or the emission maxima of catalyst **E**. On the basis of these observations, it is unlikely that base coordination to the photocatalyst greatly affects its excited-state redox properties. The observed decrease in emission intensity is probably attributed to more facile non-radiative relaxation of the excited state, in analogy to previous findings<sup>29</sup>.

Thermochemical constraints provide strong evidence against a sequential proton-transfer–electron-transfer mechanism for amidyl formation. On the basis of the difference in  $\text{p}K_a$  between the benzamide substrate and the monobasic phosphate ( $\Delta\text{p}K_a \approx 20$  in MeCN), the rate constant for proton transfer cannot exceed  $6.8 \times 10^{-11} \text{ M}^{-1} \text{ s}^{-1}$ . As such, this process would not be kinetically competitive with luminescent decay of the  $\text{Ir(III)}$  excited state ( $1/\tau_0 = 3.6 \times 10^6 \text{ s}^{-1}$ , where  $\tau_0$  is the lifetime of the excited state). Because both stepwise mechanisms can be discounted, we believe that the data presented above is most consistent with a concerted PCET mechanism of N–H bond activation. When using the catalyst **E**/dibutyl phosphate combination, the N–H scission step is calculated to be only modestly endergonic (effective BDFE of  $103 \text{ kcal mol}^{-1}$ , change in Gibbs free energy  $\Delta G = +4 \text{ kcal mol}^{-1}$ )<sup>30</sup>. Lastly, we determined that the quantum yield of the model reaction converting **1** to **2** is 0.12.

Received 26 March; accepted 25 August 2016.

Published online 12 October 2016.

1. Mayer, J. M. Understanding hydrogen atom transfer: from bond strengths to Marcus theory. *Acc. Chem. Res.* **44**, 36–46 (2011).
2. Liu, W. *et al.* Oxidative aliphatic C–H fluorination with fluoroide ion catalyzed by a manganese porphyrin. *Science* **337**, 1322–1325 (2012).
3. Jeffrey, J. L., Terrett, J. A. & MacMillan, D. W. C. O–H hydrogen bonding promotes H-atom transfer from  $\alpha$  C–H bonds for C-alkylation of alcohols. *Science* **349**, 1532–1536 (2015).
4. Chen, M. S. & White, M. C. Combined effects on selectivity in Fe-catalyzed methylene oxidation. *Science* **327**, 566–571 (2010).
5. Newhouse, T. & Baran, P. S. If C–H bonds could talk: selective C–H bond oxidation. *Angew. Chem. Int. Ed.* **50**, 3362–3374 (2011).
6. Choi, G. J. & Knowles, R. R. Catalytic alkene carboaminations enabled by oxidative proton-coupled electron transfer. *J. Am. Chem. Soc.* **137**, 9226–9229 (2015).
7. Miller, D. C., Choi, G. J., Orbe, H. S. & Knowles, R. R. Catalytic olefin hydroamination enabled by proton-coupled electron transfer. *J. Am. Chem. Soc.* **137**, 13492–13495 (2015).

8. Li, Z., Song, L. & Li, C. Silver-catalyzed radical aminofluorination of unactivated alkenes in aqueous media. *J. Am. Chem. Soc.* **135**, 4640–4643 (2013).
9. Kemper, J. & Studer, A. Stable reagents for the generation of N-centered radicals: hydroamination of norbornene. *Angew. Chem. Int. Ed.* **44**, 4914–4917 (2005).
10. Guin, J., Fröhlich, R. & Studer, A. Thiol-catalyzed stereoselective transfer hydroamination of olefins with N-aminated dihydropyridines. *Angew. Chem. Int. Ed.* **47**, 779–782 (2008).
11. Zard, S. Z. Recent progress in the generation and use of nitrogen-centered radicals. *Chem. Soc. Rev.* **37**, 1603–1618 (2008).
12. Martín, A., Pérez-Martín, I. & Suárez, E. Intramolecular hydrogen abstraction promoted by amidyl radicals. Evidence for electronic factors in the nucleophilic cyclization of ambident amides to oxocarbenium ions. *Org. Lett.* **7**, 2027–2030 (2005).
13. Chen, K., Richter, J. M. & Baran, P. S. 1,3-Diol synthesis via controlled, radical-mediated C–H functionalization. *J. Am. Chem. Soc.* **130**, 7247–7249 (2008).
14. Schmidt, V. A., Quinn, R. K., Bruscoe, A. T. & Alexanian, E. J. Site-selective aliphatic C–H bromination using N-bromoamides and visible light. *J. Am. Chem. Soc.* **136**, 14389–14392 (2014).
15. Liu, T., Mei, T.-S. & Yu, J.-Q.  $\gamma,\delta,\epsilon$ -C(sp<sup>3</sup>)–H functionalization through directed radical H-abstraction. *J. Am. Chem. Soc.* **137**, 5871–5874 (2015).
16. Horner, J. H., Musa, O. M., Bouvier, A. & Newcomb, M. Absolute kinetics of amidyl radical reactions. *J. Am. Chem. Soc.* **120**, 7738–7748 (1998).
17. Topczewski, J. J., Cabrera, P. J., Saper, N. I. & Sanford, M. S. Palladium-catalysed transannular C–H functionalization of alicyclic amines. *Nature* **531**, 220–224 (2016).
18. Zhang, F.-L., Hong, K., Li, T.-J., Park, H. & Yu, J.-Q. Functionalization of C(sp<sup>3</sup>)–H bonds using a transient directing group. *Science* **351**, 252–256 (2016).
19. Fujiwara, Y. *et al.* Practical and innate carbon–hydrogen functionalization of heterocycles. *Nature* **492**, 95–99 (2012).
20. Lin, S., Song, C.-X., Cai, G.-X., Wang, W.-H. & Shi, Z.-J. Intra/intermolecular direct allylic alkylation via Pd(II)-catalyzed allylic C–H activation. *J. Am. Chem. Soc.* **130**, 12901–12903 (2008).
21. Wolff, M. E. Cyclization of N-halogenated amines (The Hofmann–Löffler reaction). *Chem. Rev.* **63**, 55–64 (1963).
22. Bordwell, F. G., Zhang, S., Zhang, X.-M. & Liu, W.-Z. Homolytic bond dissociation enthalpies of the acidic H–A bonds caused by proximate substituents in sets of methyl ketones, carboxylic esters, and carboxamides related to changes in ground state energies. *J. Am. Chem. Soc.* **117**, 7092–7096 (1995).
23. Bordwell, F. G. & Ji, G.-Z. Effects of structural changes on acidities and homolytic bond dissociation energies of the H–N bonds in amidines, carboxamides, and thiocarboxamides. *J. Am. Chem. Soc.* **113**, 8398–8401 (1991).
24. Mayer, J. M. Proton-coupled electron transfer: a reaction chemist's view. *Annu. Rev. Phys. Chem.* **55**, 363–390 (2004).
25. Weinberg, D. R. *et al.* Proton-coupled electron transfer. *Chem. Rev.* **112**, 4016–4093 (2012).
26. Reece, S. Y. & Nocera, D. G. Proton-coupled electron transfer in biology: results from synergistic studies in natural and model systems. *Annu. Rev. Biochem.* **78**, 673–699 (2009).
27. Hanss, D., Freys, J. C., Bernardinelli, G. & Wenger, O. S. Cyclometalated iridium(III) complexes as photosensitizers for long-range electron transfer: occurrence of a Coulomb barrier. *Eur. J. Inorg. Chem.* **2009**, 4850–4859 (2009).
28. Bortolamei, N., Isse, A. A. & Gennaro, A. Estimation of standard reduction potentials of alkyl radicals involved in atom transfer radical polymerization. *Electrochim. Acta* **55**, 8312–8318 (2010).
29. Ward, W. M., Farnum, B. H., Siegler, M. & Meyer, G. J. Chloride ion-pairing with Ru(II) polypyridyl compounds in dichloromethane. *J. Phys. Chem. A* **117**, 8883–8894 (2013).
30. Warren, J. J., Tronic, T. A. & Mayer, J. M. Thermochemistry of proton-coupled electron transfer reagents and its implications. *Chem. Rev.* **110**, 6961–7001 (2010).

Supplementary Information is available in the online version of the paper.

**Acknowledgements** We thank the NIH (R01 GM113105) for financial support. R.R.K. is a Sloan Foundation Research Fellow and Amgen Young Investigator Award recipient. We also thank M. Rauch for assistance with quantum yield measurements and D. G. Oblinsky for assistance in measuring excited-state lifetimes. We thank E. Alexanian and R. Quinn (UNC Chapel Hill) for discussions. We thank T. Rovis (Colorado State University) for sharing results before publication.

**Author Contributions** R.R.K. and G.J.C. conceived of the research. G.J.C., Q.Z. and D.C.M. carried out experiments and analysed results with R.R.K. C.J.G. assisted with substrate synthesis. R.R.K. wrote the manuscript with input from all of the authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.R.K. ([rknowles@princeton.edu](mailto:rknowles@princeton.edu)).

**Reviewer Information** Nature thanks A. Studer and the other anonymous reviewer(s) for their contribution to the peer review of this work.

# Amide-directed photoredox-catalysed C–C bond formation at unactivated $sp^3$ C–H bonds

John C. K. Chu<sup>1</sup> & Tomislav Rovis<sup>1,2</sup>

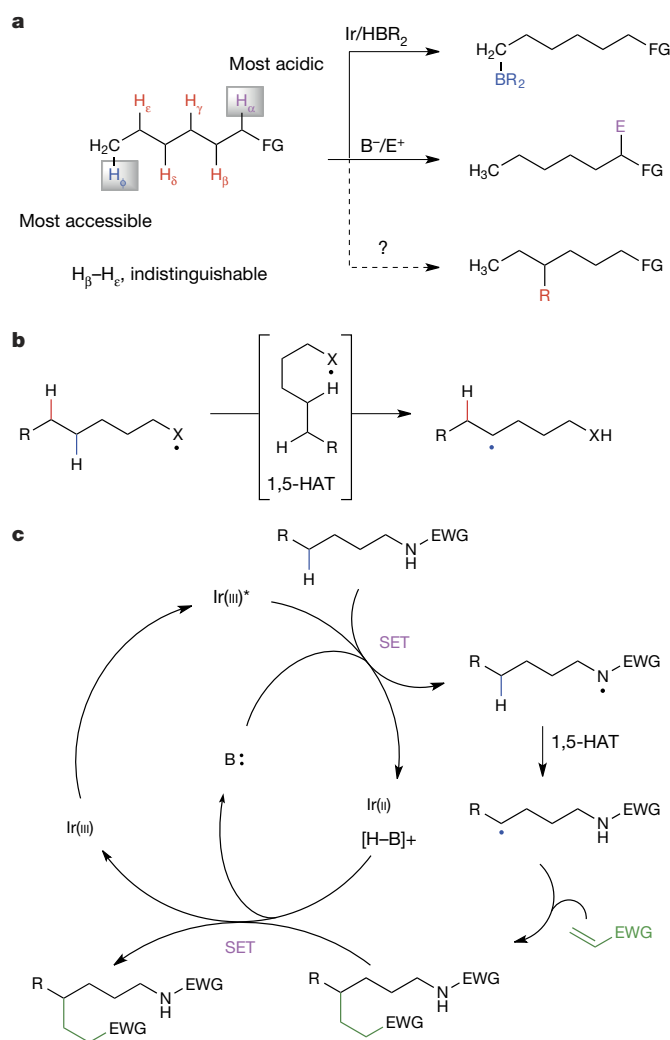
Carbon–carbon (C–C) bond formation is paramount in the synthesis of biologically relevant molecules, modern synthetic materials and commodity chemicals such as fuels and lubricants. Traditionally, the presence of a functional group is required at the site of C–C bond formation. Strategies that allow C–C bond formation at inert carbon–hydrogen (C–H) bonds enable access to molecules that would otherwise be inaccessible and the development of more efficient syntheses of complex molecules<sup>1,2</sup>. Here we report a method for the formation of C–C bonds by directed cleavage of traditionally non-reactive C–H bonds and their subsequent coupling with readily available alkenes. Our methodology allows for amide-directed selective C–C bond formation at unactivated  $sp^3$  C–H bonds in molecules that contain many such bonds that are seemingly indistinguishable. Selectivity arises through a relayed photoredox-catalysed oxidation of a nitrogen–hydrogen bond. We anticipate that our findings will serve as a starting point for functionalization at inert C–H bonds through a strategy involving hydrogen-atom transfer.

The challenges associated with targeting inert  $sp^3$  C–H bonds for C–C bond formation are their high bond strength and selectivity issues arising from their ubiquitous presence as the scaffolding of organic chemistry (Fig. 1a)<sup>3</sup>. Sterically accessible C–H bonds, which are typically termini of alkyl chains, may be activated using Rh- or Ir-based catalysts in emerging borylation chemistry<sup>4</sup>. Otherwise, inert  $sp^3$  C–H bonds are nearly impossible to distinguish in a rational and selective way, with a few exceptions<sup>5–12</sup>.

In the context of our work on Rh(III)-catalysed C–H activation<sup>13</sup>, we sought an alternative method to functionalize  $sp^3$  C–H bonds of aliphatic amines; a 1,5-hydrogen-atom transfer<sup>14,15</sup> (1,5-HAT) strategy occurred to us as a potential solution (Fig. 1b). Heteroatom radicals are known to abstract hydrogen atoms through the 1,5-HAT process in a selective way. Given the ubiquity and relevance of nitrogen in molecules of interest, we sought to generate nitrogen radicals<sup>16</sup> from highly acidified amidyl N–H bonds<sup>17</sup> using photoredox catalysis<sup>18</sup> to functionalize these distal, unactivated positions, and couple them with a subsequent C–C bond-forming step (Fig. 1c). Such a transformation would not require pre-functionalization<sup>19</sup> of the X–H bond and could comprise a net addition of the C–H bond across the alkene, an atom-economic process.

We developed reaction conditions for C–C bond formation at a tertiary C–H bond that is five bonds away from a nitrogen atom bearing a readily removable trifluoroacetyl group (Fig. 2a; see Supplementary Information for further details). This group on nitrogen acidifies the N–H bond ( $pK_a \approx 13.8$ )<sup>20</sup> sufficiently to allow deprotonation using basic  $K_3PO_4$  and to provide a driving force for 1,5-HAT. Less acidifying groups on nitrogen such as difluoroacetyl or pentafluorobenzoyl do not lead to the desired product. At this point, the excited state of the photocatalyst  $[Ir(dF-CF_3ppy)_2dtbbpy]PF_6$  (Fig. 2a) (half-cell potential  $E_{1/2}^{red}[*Ir(III)/Ir(II)] = +1.21$  V versus saturated calomel electrode in MeCN)<sup>21</sup> oxidizes the amidyl anion (peak potential  $E_p = +0.77$  V versus saturated calomel electrode in MeCN) to generate a nitrogen radical. After 1,5-HAT, a tertiary carbon radical is generated and subsequently

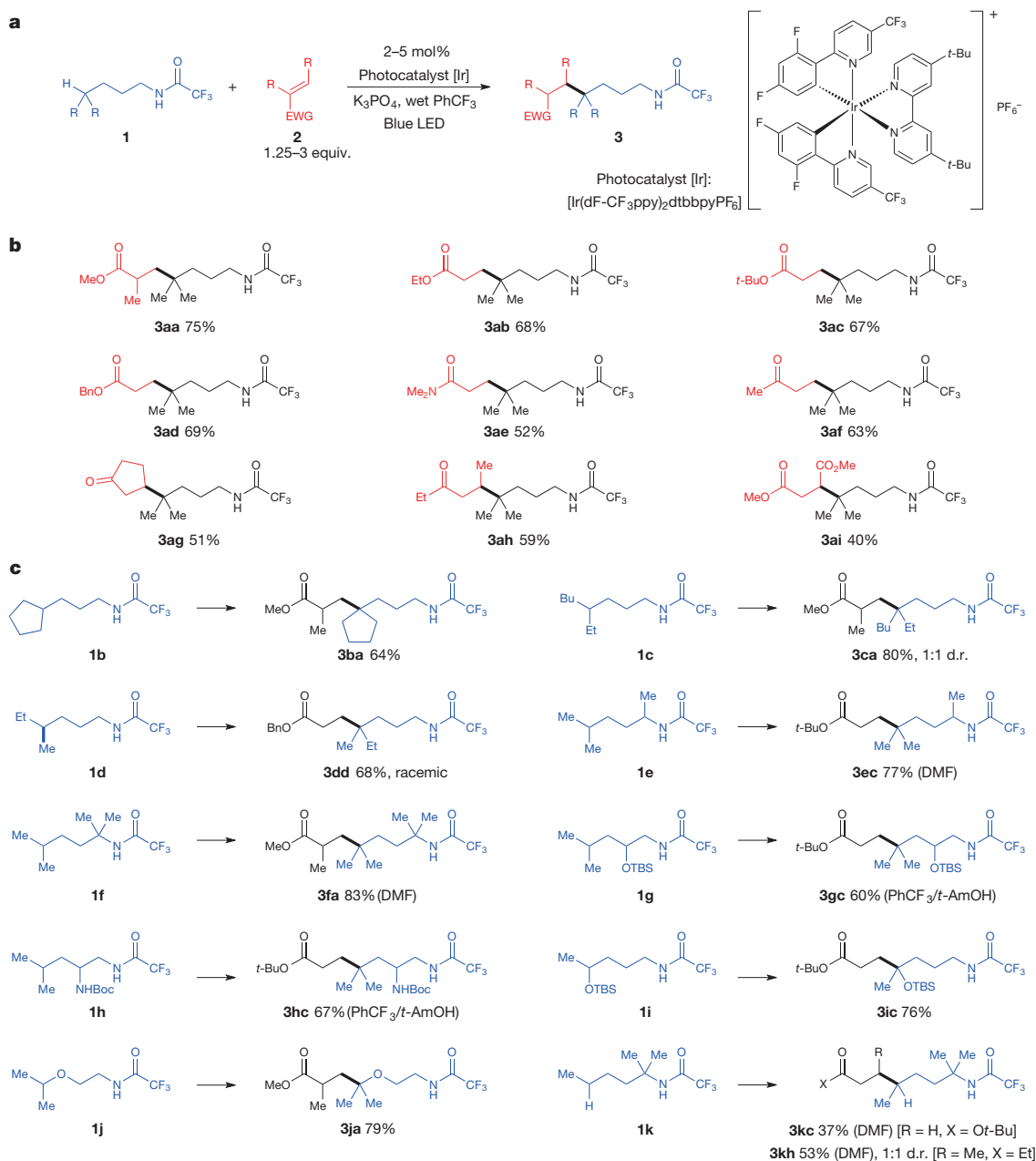
trapped with methyl methacrylate. The use of  $PhCF_3$  as a solvent and high reaction concentrations (0.4 M) were necessary to achieve high yield (Fig. 2b; see Supplementary Information for details).



**Figure 1 | Selectivity issues with C–H bond functionalization.**

**a**, Unactivated  $sp^3$  C–H bonds are similar in electronics and sterics, making them non-trivial to differentiate. **b**, Our proposed 1,5-hydrogen-atom transfer (1,5-HAT) strategy to cleave an unactivated C–H bond with a heteroatom radical. **c**, Proposed photocatalytic cycle for C–C bond formation at unactivated C–H bonds. Deprotonation of an acidified N–H bond and *in situ* oxidation delivers the nitrogen radical with subsequent 1,5-HAT and functionalization occurring. Upon trapping, the resultant radical is reduced and protonated, closing the catalytic cycle and delivering product. EWG, electron-withdrawing group; FG, functional group; SET, single electron transfer.

<sup>1</sup>Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, USA. <sup>2</sup>Department of Chemistry, Columbia University, New York, New York 10027, USA.



**Figure 2 | Photoredox-catalysed C–C bond formation at unactivated  $sp^3$  C–H bonds.** **a**, Trifluoroacetamide as the directing group for C–H functionalization. **b**, Scope of C–C bond formation with respect to the alkene. **c**, Scope of C–C bond formation with respect to the amine. Solvent is benzonitrile unless otherwise indicated. Bn, benzyl;

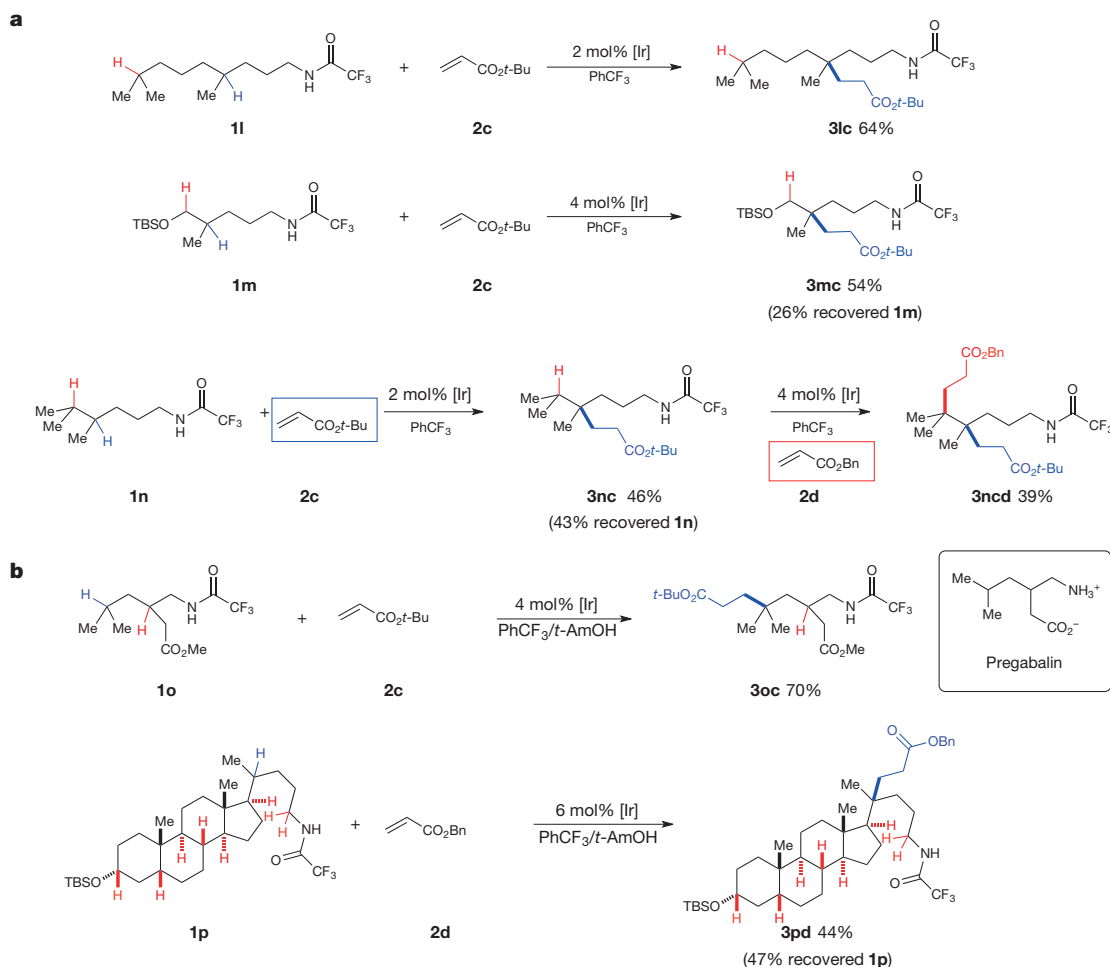
Bu, butyl; d.r., diastereomeric ratio; Et, ethyl; Me, methyl; *t*-Bu, *tert*-butyl; TBS, *tert*-butyldimethylsilyl; Boc, *tert*-butoxycarbonyl; Ph, phenyl; DMF, dimethylformamide; *t*-AmOH, *tert*-amyl alcohol; dF-CF<sub>3</sub>ppy, 2-(2,4-difluorophenyl)-5-trifluoromethylpyridine; dtbbpy, 4,4'-di-*tert*-butyl-2,2'-bipyridine.

With the optimized conditions, we investigated the variation in amine molecules with inert C–H bonds and alkenes that are applicable to the developed reaction (Fig. 2b). A range of electron-deficient alkenes were successfully trapped by the tertiary radical carbon generated through our photocatalysed 1,5-HAT methodology using trifluoroacetamide **1a**. Ethyl acrylate **2a**, an alkene devoid of an  $\alpha$ -substituent, coupled with trifluoroacetamide **1a** to give the product **3ab** in good yield. These results were mirrored with other acrylate esters such as *t*-butyl and benzyl to give products **3ac** and **3ad**, respectively. The reactive benzylic C–H bonds in benzyl acrylate, or in the corresponding product **3ad**, remain intact under the reaction conditions. *N,N*-dimethyl acrylamide **2e**, which has a lower reactivity compared to acrylates, couples with trifluoroacetamide **1a** in reasonable yield. Coupling with vinyl ketones was possible, giving products **3af**,

**3ag** and **3ah** in good yields. The use of methyl vinyl ketone requires lower reaction concentrations and higher catalyst loading to disfavour a competitive aza-Michael reaction pathway, which is formed in >50% conversion in the absence of the photoredox catalyst. Less reactive  $\beta$ -substituted vinyl ketones require a higher alkene concentration. Product **3ai** was obtained in moderate yield using dimethyl maleate.

Variation of the steric environment around the inert C–H bonds had little effect on the outcome of the reaction; products **3ba**, **3ca** and **3dd** were all obtained in good yield (Fig. 2c). Subjection of enantiomerically pure alkane possessing a stereogenic C–H bond results in formation of racemic product **3dd**, consistent with our proposed mechanism. Substrates that bear  $\alpha$ -substituents to nitrogen were also competent in the reaction, delivering products **3ec** and **3fa** in good yields. In these





**Figure 3 | Regioselective functionalization of  $sp^3$  C–H bonds.** **a**, The presence of a remote tertiary C–H bond does not affect the outcome of the reaction; good selectivity is obtained even with potential competition

cases, optimal yields were obtained by switching from  $\text{PhCF}_3$  to the more polar solvent DMF. Heteroatom functionality including a TBS-protected oxygen and a Boc-protected amine can be incorporated into the trifluoroacetamide substrate to give products **3gc** and **3hc** in good yields. Importantly, the presence of a Boc-protected primary amine in substrate **1h** does not affect the outcome of the reaction, presumably because the Boc–N–H bond is not acidic enough to be deprotonated by the  $\text{K}_3\text{PO}_4$ . Substrates **1i** and **1j**, with oxygen atoms adjacent to the C–H bond, participate well in the reaction to give products **3ic** and **3ja**. Methylene C–H bonds may also be functionalized using this approach, as demonstrated with substrate **1k** to deliver product **3kc**, albeit in slightly lower yield. Increased yields are observed using a  $\beta$ -substituted Michael acceptor, delivering products such as **3kh**.

Substrates that contain more than one tertiary C–H bond were explored in the reaction (Fig. 3a). Trifluoroacetamide **1l**, which contains two tertiary C–H bonds, can be selectively functionalized at the position proximal to the N–H bond. Functionalization of the remote tertiary C–H bond is not observed. This outcome supports our hypothesis that an intramolecular 1,5-HAT process is operating. When there is potential for a competitive 1,6-HAT<sup>22</sup>, as in substrates **1m** and **1n**, 1,5-HAT is favoured over 1,6-HAT, resulting in products **3mc** and **3nc**. Although 1,5-HAT occurs in preference to 1,6-HAT, products arriving from sequential 1,5-HAT and then 1,6-HAT are observed when reactions are run to completion. Re-subjecting product **3nc** to the reaction conditions, but with a different alkene coupling partner, results in product **3ncd** in moderate yield. This reaction sequence allows for the coupling of two distinct alkene partners at the two tertiary C–H bonds

between 1,5- and 1,6-hydrogen-atom transfer. **b**, Applications to medicinally relevant molecules. Bn, benzyl; Me, methyl; *t*-Bu, *tert*-butyl; TBS, *tert*-butyldimethylsilyl; *t*-AmOH, *tert*-amyl alcohol; Ph, phenyl.

in substrate **1n** with control over regioselectivity to give a molecule containing two contiguous all-carbon quaternary centres. Medicinally relevant molecules such as Pregabalin<sup>23</sup> can also be functionalized (Fig. 3b). Lastly, steroid-derived trifluoroacetamide **1p**, a molecule containing six distinct tertiary C–H bonds and three C–H bonds adjacent to heteroatoms, can be selectively functionalized at a single position, determined by its proximity to the N–H bond.

Our mechanistic studies (see Supplementary Information for a full discussion) suggest that a stepwise deprotonation/oxidation event is at least partly responsible for the generation of the nitrogen radical. Both cyclic voltammetry and Stern–Volmer studies show that the trifluoroacetamide conjugate base quenches the excited Ir photocatalyst. The similarity between the  $\text{p}K_a$  values for trifluoroacetamide and the conjugate acid of  $\text{K}_3\text{PO}_4$  under these biphasic conditions means that appreciable amounts of the trifluoroacetamide conjugate base are present in solution. The observation of aza-Michael product with methyl vinyl ketone supports the assertion that the amidyl anion is formed in appreciable amounts. Furthermore, weaker electron-withdrawing groups on nitrogen such as difluoroacetyl do not lead to either the desired product or the aza-Michael adduct, which suggests that a negative charge on nitrogen is necessary to obtain the desired reactivity. Therefore, the bulk of the evidence supports the idea that the stepwise pathway is operative. On the other hand, the use of  $\text{Cs}_2\text{CO}_3$ , a weaker base ( $\text{p}K_a = 10.3$ ), also provides the desired product, albeit in lower yield. This raises the possibility that the proton-coupled electron-transfer pathway<sup>24,25</sup> can be operative in tandem with the stepwise mechanism.

In summary, we have developed a photoredox-catalysed amide-directed selective  $sp^3$  C–H bond functionalization with concomitant C–C bond formation. The proposed mechanism features the formation of a nitrogen-centred radical by photoredox-mediated oxidation of an amidyl anion formed *in situ*, and subsequent 1,5-HAT to relay the reactive radical species to a specific carbon atom. The subsequent radical can be trapped with several electron-deficient alkenes. This finding addresses one of the shortcomings in aliphatic C–H bond functionalization.

Received 25 March; accepted 25 August 2016.

Published online 12 October 2016.

- Gutekunst, W. R. & Baran, P. S. C–H functionalization logic in total synthesis. *Chem. Soc. Rev.* **40**, 1976–1991 (2011).
- Yamaguchi, J., Yamaguchi, A. D. & Itami, K. C–H bond functionalization: emerging synthetic tools for natural products and pharmaceuticals. *Angew. Chem. Int. Ed.* **51**, 8960–9009 (2012).
- Newhouse, T. & Baran, P. S. If C–H bonds could talk: selective C–H bond oxidation. *Angew. Chem. Int. Ed.* **50**, 3362–3374 (2011).
- Lawrence, J. D., Takahashi, M., Bae, C. & Hartwig, J. F. Regiospecific functionalization of methyl C–H bonds of alkyl groups in reagents with heteroatom functionality. *J. Am. Chem. Soc.* **126**, 15334–15335 (2004).
- Shabashov, D. & Daugulis, O. Auxiliary-assisted palladium-catalyzed arylation and alkylation of  $sp^2$  and  $sp^3$  carbon–hydrogen bonds. *J. Am. Chem. Soc.* **132**, 3965–3972 (2010).
- He, J. *et al.* Ligand-controlled  $C(sp^3)$ –H arylation and olefination in synthesis of unnatural chiral  $\alpha$ -amino acids. *Science* **343**, 1216–1220 (2014).
- Jiang, H., He, J., Liu, T. & Yu, J.-Q. Ligand enabled  $\gamma$ - $C(sp^3)$ –H olefination of amines: en route to pyrrolidines. *J. Am. Chem. Soc.* **138**, 2055–2059 (2016).
- Zhang, F.-L., Hong, K., Li, T.-J., Park, H. & Yu, J.-Q. Functionalization of  $C(sp^3)$ –H bonds using a transient directing group. *Science* **351**, 252–256 (2016).
- Topczewski, J. J., Cabrera, P. J., Saper, N. I. & Sanford, M. S. Palladium-catalysed transannular C–H functionalization of alicyclic amines. *Nature* **531**, 220–224 (2016).
- Chen, M. S. & White, M. C. Combined effects on selectivity in Fe-catalyzed methylene oxidation. *Science* **327**, 566–571 (2010).
- Schmidt, V. A., Quinn, R. K., Brusoe, A. T. & Alexanian, E. J. Site-selective aliphatic C–H bromination using N-bromoamides and visible light. *J. Am. Chem. Soc.* **136**, 14389–14392 (2014).
- McNally, A., Haffemayer, B., Collins, B. S. L. & Gaunt, M. J. Palladium-catalysed C–H activation of aliphatic amines to give strained nitrogen heterocycles. *Nature* **510**, 129–133 (2014).
- Archambeau, A. & Rovis, T. Rhodium(III)-catalyzed allylic  $C(sp^3)$ –H activation of alkenyl sulfonamides: unexpected formation of azabicycles. *Angew. Chem. Int. Ed.* **54**, 13337–13340 (2015).
- Robertson, J., Pillai, J. & Lush, R. K. Radical translocation reactions in synthesis. *Chem. Soc. Rev.* **30**, 94–103 (2001).
- Cekovic, Z. Reactions of carbon radicals generated by 1,5-transposition of reactive centers. *J. Serb. Chem. Soc.* **70**, 287–318 (2005).
- Zard, S. Z. Recent progress in the generation and use of nitrogen-centered radicals. *Chem. Soc. Rev.* **37**, 1603–1618 (2008).
- Nikishin, G. I., Troyansky, E. I. & Lazareva, M. I. Regiospecific oxidative cyclization of N-methylsulfonamides into pyrrolidines. *Tetrahedron* **41**, 4279–4288 (1985).
- Noble, A. & MacMillan, D. W. C. Photoredox  $\alpha$ -vinylolation of  $\alpha$ -amino acids and N-aryl amines. *J. Am. Chem. Soc.* **136**, 11602–11605 (2014).
- Zhang, J., Li, Y., Zhang, F., Hu, C. & Chen, Y. Generation of alkoxy radicals by photoredox catalysis enables selective  $C(sp^3)$ –H functionalization under mild reaction conditions. *Angew. Chem. Int. Ed.* **55**, 1872–1875 (2016).
- Li, J. J. *Named Reactions for Functional Group Transformation* 423–437 (Wiley, 2007).
- Lowry, M. S. *et al.* Single-layer electroluminescent devices and photoinduced hydrogen production from an ionic iridium(III) complex. *Chem. Mater.* **17**, 5712–5719 (2005).
- Nechab, M., Mondal, S. & Bertrand, M. P. 1,*n*-Hydrogen-atom transfer (HAT) reactions in which  $n \neq 5$  an updated inventory. *Chem. Eur. J.* **20**, 16034–16059 (2014).
- Tassone, D. M., Boyce, E., Guyer, J. & Nuzum, D. Pregabalin: a novel  $\gamma$ -aminobutyric acid analogue in the treatment of neuropathic pain, partial-onset seizures, and anxiety disorders. *Clin. Ther.* **29**, 26–48 (2007).
- Choi, G. J. & Knowles, R. R. Catalytic alkene carboaminations enabled by oxidative proton-coupled electron transfer. *J. Am. Chem. Soc.* **137**, 9226–9229 (2015).
- Miller, D. C., Choi, G. J., Orbe, H. S. & Knowles, R. R. Catalytic olefin hydroamidation enabled by proton-coupled electron transfer. *J. Am. Chem. Soc.* **137**, 13492–13495 (2015).

Supplementary Information is available in the online version of the paper.

**Acknowledgements** We thank NIGMS (GM80442) for support. J.C.K.C. thanks the Croucher Foundation (Hong Kong) for support. We thank M. Burns (CSU) for technical assistance. We thank R. R. Knowles (Princeton University) for sharing results before publication.

**Author Contributions** T.R. and J.C.K.C. conceived the concept and prepared the manuscript. T.R. directed the investigation. J.C.K.C. developed and studied the reaction.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.R. (tr2504@columbia.edu).

**Reviewer Information** Nature thanks A. Studer and the other anonymous reviewer(s) for their contribution to the peer review of this work.

# Fluvial sediment supply to a mega-delta reduced by shifting tropical-cyclone activity

Stephen E. Darby<sup>1</sup>, Christopher R. Hackney<sup>1</sup>, Julian Leyland<sup>1</sup>, Matti Kummu<sup>2</sup>, Hannu Lauri<sup>3</sup>, Daniel R. Parsons<sup>4</sup>, James L. Best<sup>5</sup>, Andrew P. Nicholas<sup>6</sup> & Rolf Aalto<sup>6</sup>

The world's rivers deliver 19 billion tonnes of sediment to the coastal zone annually<sup>1</sup>, with a considerable fraction being sequestered in large deltas, home to over 500 million people. Most (more than 70 per cent) large deltas are under threat from a combination of rising sea levels, ground surface subsidence and anthropogenic sediment trapping<sup>2,3</sup>, and a sustainable supply of fluvial sediment is therefore critical to prevent deltas being 'drowned' by rising relative sea levels<sup>2–4</sup>. Here we combine suspended sediment load data from the Mekong River with hydrological model simulations to isolate the role of tropical cyclones in transmitting suspended sediment to one of the world's great deltas. We demonstrate that spatial variations in the Mekong's suspended sediment load are correlated ( $r = 0.765$ ,  $P < 0.1$ ) with observed variations in tropical-cyclone climatology, and that a substantial portion (32 per cent) of the suspended sediment load reaching the delta is delivered by runoff generated by rainfall associated with tropical cyclones. Furthermore, we estimate that the suspended load to the delta has declined by  $52.6 \pm 10.2$  megatonnes over recent years (1981–2005), of which  $33.0 \pm 7.1$  megatonnes is due to a shift in tropical-cyclone climatology. Consequently, tropical cyclones have a key role in controlling the magnitude of, and variability in, transmission of suspended sediment to the coast. It is likely that anthropogenic sediment trapping in upstream reservoirs is a dominant factor in explaining past<sup>5–7</sup>, and anticipating future<sup>8,9</sup>, declines in suspended sediment loads reaching the world's major deltas. However, our study shows that changes in tropical-cyclone climatology affect trends in fluvial suspended sediment loads and thus are also key to fully assessing the risk posed to vulnerable coastal systems.

The world's largest rivers contribute a disproportionately large fraction (Extended Data Table 1) of the terrestrial sediment flux, which has both created, and is critical in sustaining, their great deltas. Moreover, river-borne sediments are a key vector for carbon and nutrients, thereby playing a vital part in global biogeochemical cycles<sup>10,11</sup>. However, a majority (>70%) of large deltas are now recognized as being under severe threat from rising relative sea levels<sup>2,3</sup>, in part owing to reported anthropogenically driven reductions in sediment loads<sup>5–7</sup>. Many large rivers are located in tropical regions (Extended Data Fig. 1) that exhibit highly seasonal flow regimes affected by tropical cyclones (TCs). The potential destructive or constructive impacts of TCs that directly strike deltas are well established<sup>12,13</sup>. However, when they strike further upstream, TCs deliver much higher than normal levels of rainfall, effectively triggering landslides and mobilizing sediments into the river network, thereby generating very high instantaneous sediment loads<sup>14–16</sup>. Such high-sediment loads could compensate for the potential destructive effects of TCs striking deltas proper, but, notwithstanding some previous studies in smaller drainage basins<sup>17,18</sup>, the role of TCs in driving sediment delivery to the lowlands and coast remains unclear. As noted, this is particularly the case for large rivers that carry much of the terrestrial sediment flux because these rivers are, in their

mid to lower reaches, typically bound by massive floodplains that can sequester considerable volumes of suspended sediment into storage during floods<sup>19</sup>. Here we address this uncertainty by quantifying the role of TCs in driving suspended sediment loads through an exemplar mega-river, the Mekong.

Draining the Tibetan Plateau and the Annamite Mountains bordering Laos and Vietnam (Fig. 1), and with the monsoonal climate generating intense rainfall, the Mekong basin (795,000 km<sup>2</sup>) generates fluxes of water (450 km<sup>3</sup> yr<sup>−1</sup>) (ref. 20) and sediment (~160 Mt yr<sup>−1</sup>, but see later)<sup>21</sup> that rank tenth and ninth, respectively, among the world's great rivers<sup>1</sup>. The Mekong is therefore similar to other major rivers (for example, Ganges–Brahmaputra, Yangtze and Mississippi) that transmit globally important sediment loads and that are influenced in their mid to lower courses by TCs. Similar to these other rivers, the sediments of the Mekong River have resulted in the formation of a large delta, with much contemporary debate on the extent to which declining sediment loads may in the future increase the vulnerability of the Mekong Delta to rising sea levels<sup>8,9,22</sup>.

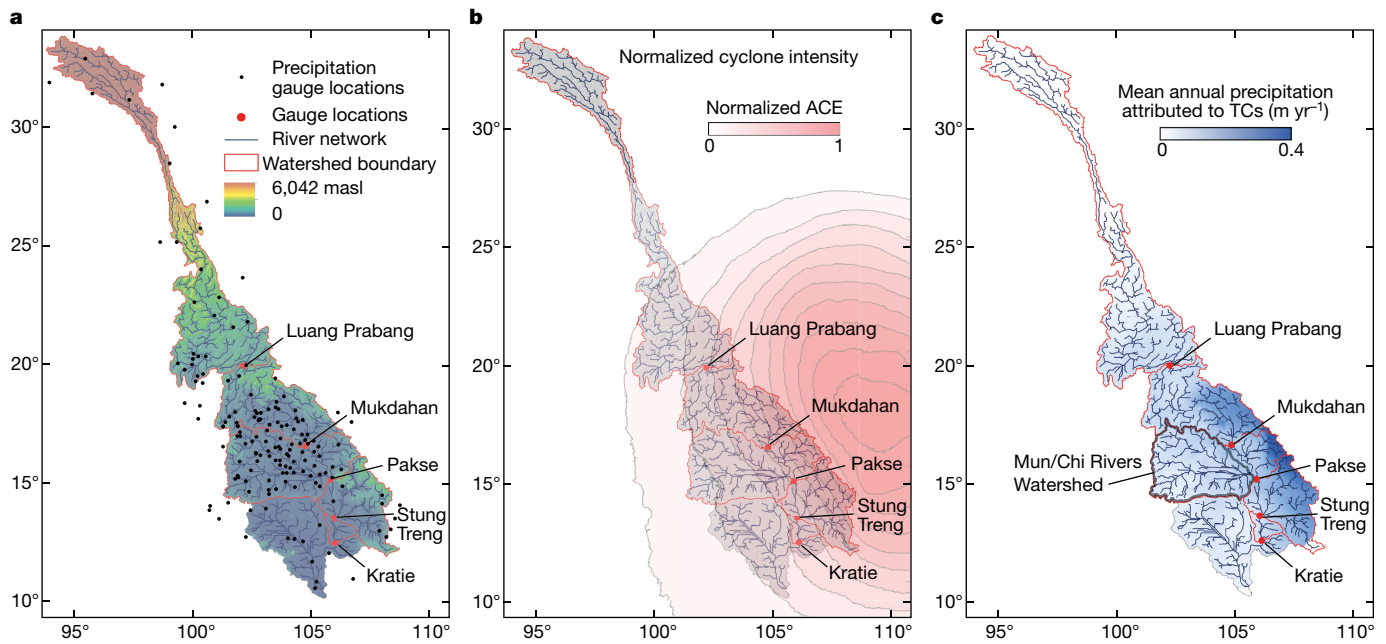
To quantify the influence of TCs on the suspended sediment transport regime, we determined temporal (25 years) and spatial (1,400 km study reach) variations in suspended solids loads throughout the Lower Mekong River (see Methods). Specifically, we first used a distributed hydrological model, forced with two climate scenarios, one with and the other without observed TCs, to simulate water discharges at five river gauging stations (see Methods for model details and Fig. 1 for gauging station locations): Luang Prabang in Laos, Mukdahan in Thailand, Pakse in Laos, and Stung Treng and Kratie, both in Cambodia. Importantly, these five river gauging stations are situated on an environmental gradient that spans regions that are weakly (Luang Prabang) to moderately (Mukdahan and Pakse) to strongly (Stung Treng and Kratie) affected by TCs (Fig. 1b, c). We then analysed archival measurements of suspended solids concentration, collected by the respective national hydrological agencies, to construct new suspended sediment rating curves—statistical functions linking the rate of suspended sediment transport to water discharge—for the five stations (see Methods and Extended Data Fig. 2). These rating curves were then used with the model-simulated water discharges to compute suspended solids loads and to apportion these loads into TC-driven components ( $Q_{s\_TC}$ ) using:

$$Q_{s\_TC} = Q_s \left( \frac{Q_{sim\_TC}}{Q_{sim}} \right) \quad (1)$$

where  $Q_s$  is the total suspended solids load as computed using the sediment rating curves with the total simulated flow discharge,  $Q_{sim}$  (that is, the flow discharge for the baseline scenario with the observed climatology including TCs), and  $Q_{sim\_TC}$  is the simulated flow discharge attributable to TCs. The quantity  $Q_{sim\_TC}$  in equation (1) is determined by differencing the flow discharges computed in the two scenarios with ( $Q_{sim}$ ) and without ( $Q_{no\_TC}$ ) TCs, such that  $Q_{sim\_TC} = Q_{sim} - Q_{no\_TC}$ .

<sup>1</sup>Geography and Environment, University of Southampton, Southampton SO17 1BJ, UK. <sup>2</sup>Water and Development Research Group, Aalto University, 02150 Espoo, Finland. <sup>3</sup>EIA Finland Ltd, Sinimäentie 10B, 02630 Espoo, Finland. <sup>4</sup>Department of Geography, Environment and Earth Sciences, University of Hull, Hull HU6 7RX, UK. <sup>5</sup>Departments of Geology, Geography & GIS, Mechanical Science and Engineering and Ven Te Chow Hydrosystems Laboratory, University of Illinois, Champaign, Illinois 61820, USA. <sup>6</sup>Department of Geography, University of Exeter, Exeter EX4 4RJ, UK.





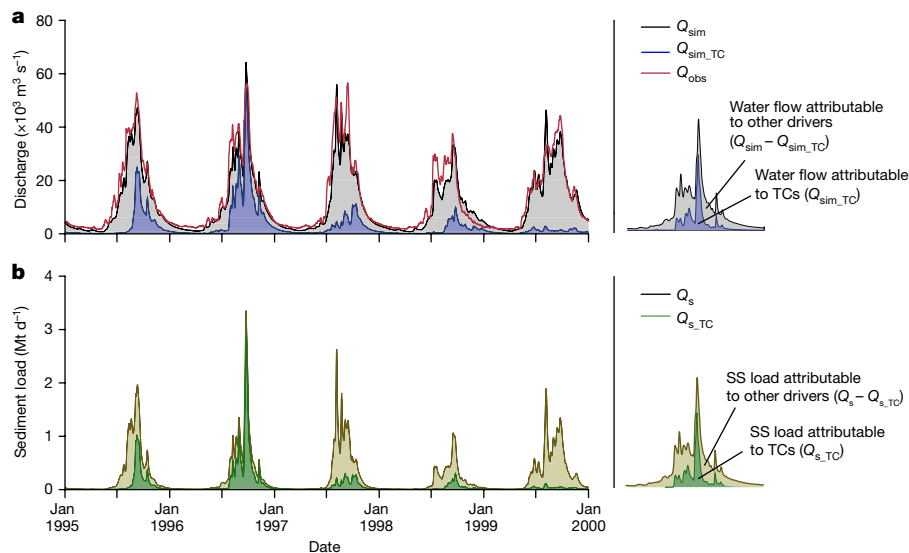
**Figure 1 | The gauging network of the Mekong River.** **a**, Gauging stations at Luang Prabang (drainage area,  $A = 323,600 \text{ km}^2$ ), Mukdahan ( $A = 464,100 \text{ km}^2$ ), Pakse ( $A = 632,600 \text{ km}^2$ ), Stung Treng ( $A = 722,300 \text{ km}^2$ ) and Kratie ( $A = 734,200 \text{ km}^2$ ), also showing the basin topography and network of meteorological stations. masl, metres

above sea level. **b**, TC climatology (1981–2005) as represented by the normalized ACE metric. **c**, Mean annual rainfall (1981–2005) from TCs. The highlighted Mun and Chi Rivers drain a region weakly affected by cyclones.

The hydrological model predicts water discharges that closely match historical records (as an example we show data for Kratie in Fig. 2a, but results for all the other stations are shown in Extended Data Fig. 3). Notable peaks and troughs in the total simulated flow discharge ( $Q_{\text{sim}}$ ) and the flow discharge attributable to TCs ( $Q_{\text{sim,TC}}$ ) are evident. These variable flows force marked fluctuations in simulated instantaneous suspended sediment loads, but notably there are multiple TC-forced suspended sediment transport events in most years (as indicated by the peaks in Fig. 2b). Integrating over the 25-year study period then yields estimates of mean annual suspended sediment load (Extended Data Table 2). Our estimate for Kratie ( $87.4 \pm 28.7 \text{ Mt yr}^{-1}$ ), the station closest to the apex of the Mekong Delta, falls within the lower limit

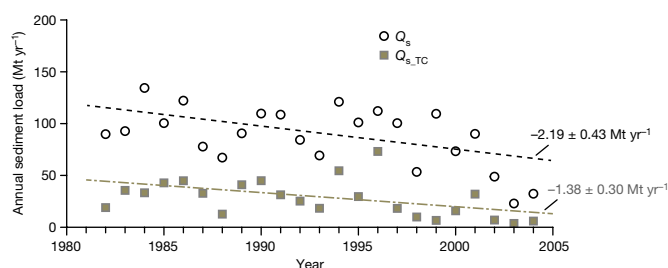
of the range ( $\sim 81\text{--}111 \text{ Mt yr}^{-1}$ ) of recent estimates<sup>23,24</sup>, although it is substantially less than the  $\sim 150\text{--}170 \text{ Mt yr}^{-1}$  cited by older studies<sup>1,20,25</sup> based on less reliable data sets.

Importantly, our results illustrate the extent to which the modest (at annual timescales) rainfall totals associated with TCs nevertheless effectively generate runoff and suspended sediment transport. During 1981–2005, TCs delivered only between 1.8% (above Luang Prabang) and 4.7% (above Kratie) of annual rainfall, but generated between 13.7% (Luang Prabang) and 28.8% (Kratie) of annual runoff. The proportion of the mean annual suspended sediment load forced by TC-associated runoff is greater still, varying between 15.2% (Luang Prabang) and 31.7% (Kratie) (Extended Data Table 2).



**Figure 2 | Daily flow discharge and suspended solids load at Kratie from 1 January 1995 to 31 December 1999.** **a**, Daily simulated ( $Q_{\text{sim}}$ ) and observed ( $Q_{\text{obs}}$ ) water flows, along with the daily water flows attributable to tropical cyclones ( $Q_{\text{sim,TC}}$ ). **b**, Daily total suspended solids (SS) load

( $Q_s$ ; in Mt per day) and daily suspended solids load attributable to TCs ( $Q_{s,TC}$ ; also in Mt per day). Note that the period 1995 to 1999 encompasses the years during the 1981–2005 study period that are the most (1996) and least (1999) strongly affected by TCs. Jan, January.



**Figure 3 | Time series of annual suspended solids loads at Kratie from 1982 to 2004.** The total annual suspended solids load ( $Q_s$ ; in Mt per year) and suspended solids load attributable to TCs ( $Q_{s\_TC}$ ; also in Mt per year) are shown with significant ( $P < 0.05$ ) trends as identified by Mann-Kendall analysis indicated by the dashed lines. The numerical value of the time rate of change of annual suspended solids load (with error) is also indicated for each of the trend lines.

There are two reasons for this amplification effect. First, TC-derived rainfall is strongly seasonal, falling largely during, or just after, the monsoon months, when catchments are pre-wetted; consequently TC-associated rainfall is very effective in generating runoff<sup>26</sup>. Second, the sediment rating functions linking suspended sediment flux and water discharge possess exponents with values exceeding unity (Extended Data Fig. 2), meaning that the peak flows generated by TCs promote very high instantaneous suspended sediment fluxes. Therefore, suspended sediment transport associated with TCs contributes substantially to mean annual loads, with the former correlating well ( $r = 0.765$ ,  $P = 0.099$ ) with the time-averaged TC climatology as represented by the 1981–2005 accumulated cyclone energy (ACE; Extended Data Table 2).

Temporal trends in annual suspended sediment load ( $Q_s$ ), and the component of that load associated with TCs ( $Q_{s\_TC}$ ), during 1981–2005 are shown for Kratie in Fig. 3 (results for all the other stations are shown in Extended Data Fig. 4). Nonparametric Mann-Kendall tests (see Methods) reveal that there have been declines in both  $Q_s$  and  $Q_{s\_TC}$  at three (Mukdahan, Stung Treng and Kratie) of the four stations that are either moderately or strongly influenced by TCs (the exception is Pakse, as discussed later). As expected, the station that is only weakly affected by TCs (Luang Prabang, Extended Data Fig. 4a) does not exhibit any significant trends in  $Q_s$  or  $Q_{s\_TC}$  that are not artefacts of the response of this station to upstream damming. Importantly, recent historical declines in  $Q_s$  at Mukdahan, Stung Treng and Kratie (Extended Data Fig. 4 and Fig. 3) are driven to a large extent by declines in the suspended sediment load attributable to TCs ( $Q_{s\_TC}$ ). Specifically, at Mukdahan, 62% of the 21.4 Mt decline in  $Q_s$  between 1981 and 2005 is attributable to reducing  $Q_{s\_TC}$  (Extended Data Fig. 4b). At the Cambodian stations, 44% (Stung Treng; Extended Data Fig. 4d) and 61% (Kratie; Fig. 3) of the declines in  $Q_s$  are attributable to reducing  $Q_{s\_TC}$ . Thus, the response of  $Q_s$  over time is intimately tied to the extent to which upstream catchments receive TC-derived rainfall (Extended Data Fig. 5).

As noted earlier, Pakse is exceptional in that it is moderately influenced by TCs (4.1% of annual rainfall is associated with TCs), but TC-driven runoff (8.4%) and suspended sediment loads (9.3%) are both anomalously low compared to Mukdahan, Stung Treng and Kratie (Extended Data Table 2). However, TC-associated rainfall is less hydrologically effective at Pakse because flows there are also strongly influenced by inflows from a major west bank tributary system, the Mun/Chi, that joins immediately upstream of the gauge and which drains a region that is only mildly influenced by TCs (Fig. 1). Additionally, the exponent in the suspended sediment rating curve at Pakse is much less than those at Stung Treng and Kratie (Extended Data Fig. 2), meaning the higher flows associated with TCs generate comparatively lower instantaneous suspended sediment transport rates.

Our results demonstrate that TCs are effective in transmitting suspended sediment load through the lowlands of large rivers, a finding

that has profound implications. A substantial portion (~40%–50%) (ref. 27) of the suspended sediment load of the Mekong River is deposited in its delta, home to 20 million people and the rice basket of Southeast Asia<sup>22,28</sup>. Major concerns have been raised regarding the scale of recent and projected future reductions in the sediment load reaching the delta, as a result of sand mining<sup>22,29</sup> and upstream damming<sup>8,9,22,23</sup>. However, our study reveals that during the period 1981–2005, the Mekong at Kratie is estimated to have experienced a cumulative loss of  $33.0 \pm 7.1$  Mt of its suspended sediment load (Fig. 3) as a result of changes in precipitation delivered by TCs crossing the Mekong basin (Extended Data Fig. 5).

Limitations in the observational data make it challenging to contextualize fully the 1981–2005 trends in TC climatology that are the focus of this paper within the longer-term historical record (Extended Data Fig. 6). Nevertheless, our key finding, namely that changes in TC climatology represent an important, but previously neglected, driver of suspended sediment transmission through the Mekong River, remains robust. Furthermore, high-resolution climate models indicate that although the number and intensity of TCs tracking across the South China Sea will probably increase under future anthropogenic climate change, their track locations will shift eastwards and away from the Indochina peninsula, leading to net reductions in ACE over the Mekong basin<sup>30</sup>. If these projected reductions in ACE are correct, TC-driven suspended sediment delivery to the Mekong Delta will decline still further, exacerbating projected declines in sediment loads due to damming<sup>8,9</sup> and sand mining<sup>29</sup> and placing the delta at even greater risk. Although our data focus on the suspended sediment load, the delivery of bedload sediment, which is important in the construction, or restoration, of deltas<sup>31</sup>, would also be lessened by a reduction in TC-associated sedimentation. Furthermore, other large rivers that transport a considerable proportion of the global sediment flux are also affected by TCs (Extended Data Table 1). Our study indicates that their deltas may also be much more affected by, and vulnerable to, changes in TC climatology than assumed in current assessments (which tend to focus on the direct effects of cyclone strikes within deltas, rather than the upstream impacts that we have focused on) of the impacts of future environmental change.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 3 February; accepted 22 August 2016.**

**Published online 19 October 2016.**

1. Milliman, J. D. & Farnsworth, K. L. *River Discharge to the Coastal Ocean: A Global Synthesis* (Cambridge Univ. Press, 2011).
2. Ericson, J. P., Vörösmarty, C. J., Dingman, S. L., Ward, L. G. & Meybeck, M. Effective sea-level rise and deltas: causes of change and human dimension implications. *Global Planet. Change* **50**, 63–82 (2006).
3. Syvitski, J. P. M. *et al.* Sinking deltas due to human activities. *Nat. Geosci.* **2**, 681–686 (2009).
4. Darby, S. E., Dunn, F. E., Nicholls, R. J., Rahman, M. & Riddy, L. A first look at the influence of anthropogenic climate change on the future delivery of fluvial sediment to the Ganges–Brahmaputra–Meghna delta. *Environ. Sci. Process. Impacts* **17**, 1587–1600 (2015).
5. Vörösmarty, C. J. *et al.* Anthropogenic sediment retention: major global impact from registered river impoundments. *Global Planet. Change* **39**, 169–190 (2003).
6. Walling, D. E. & Fang, D. Recent trends in the suspended sediment loads of the world's rivers. *Global Planet. Change* **39**, 111–126 (2003).
7. Wang, S. *et al.* Reduced sediment transport in the Yellow River due to anthropogenic changes. *Nat. Geosci.* **9**, 38–41 (2016).
8. Kumm, M. J., Wang, J. J., Lu, X. X. & Varis, O. Basin-wide sediment trapping efficiency of emerging reservoirs along the Mekong. *Geomorphology* **119**, 181–197 (2010).
9. Kondolf, G. M., Rubin, Z. K. & Minear, J. T. Dams on the Mekong: cumulative sediment starvation. *Wat. Resour. Res.* **50**, 5158–5169 (2014).
10. Richey, J. E., Brock, J. T., Naiman, R. J., Wissmar, R. C. & Stallard, R. F. Organic carbon: oxidation and transport in the Amazon River. *Science* **207**, 1348–1351 (1980).
11. Aufdenkampe, A. K. *et al.* Riverine coupling of biogeochemical cycles between land, oceans, and atmosphere. *Front. Ecol. Environ.* **9**, 53–60 (2011).
12. Day, J. W. Jr *et al.* Restoration of the Mississippi Delta: lessons from hurricanes Katrina and Rita. *Science* **315**, 1679–1684 (2007).

13. Duc, D. M., Nhuan, M. T. & Ngoi, C. V. An analysis of coastal erosion in the tropical rapid accretion delta of the Red River, Vietnam. *J. Asian Earth Sci.* **43**, 98–109 (2012).
  14. Milliman, J. D. & Kao, S. J. Hyperpycnal discharge of fluvial sediment to the ocean: impact of super-typhoon Herb (1996) on Taiwanese rivers. *J. Geol.* **113**, 503–516 (2005).
  15. Dadson, S. J. *et al.* Links between erosion, runoff variability and seismicity in the Taiwan orogen. *Nature* **426**, 648–651 (2003).
  16. Hilton, R. G. *et al.* Tropical-cyclone driven erosion of the terrestrial biosphere from mountains. *Nat. Geosci.* **1**, 759–762 (2008).
  17. Terry, J. P., Garimella, S. & Kostaschuk, R. A. Rates of floodplain accretion in a tropical island river system impacted by cyclones and large floods. *Geomorphology* **42**, 171–182 (2002).
  18. Amos, K. J. *et al.* Supply limited sediment transport in a high-discharge event of the tropical Burdekin River, North Queensland, Australia. *Sedimentology* **51**, 145–162 (2004).
  19. Aalto, R. *et al.* Episodic sediment accumulation on Amazonian flood plains influenced by El Niño/Southern Oscillation. *Nature* **425**, 493–497 (2003).
  20. Mekong River Commission. *Overview of the Hydrology of the Mekong River Basin* (Mekong River Commission, 2005).
  21. Milliman, J. D. & Meade, R. H. World-wide delivery of river sediment to the oceans. *J. Geol.* **91**, 1–21 (1983).
  22. Anthony, E. J. *et al.* Linking rapid erosion of the Mekong River delta to human activities. *Sci. Rep.* **5**, 14745 (2015).
  23. Kumm, M. & Varis, O. Sediment-related impacts due to upstream reservoir trapping on the Lower Mekong River. *Geomorphology* **85**, 275–293 (2007).
  24. Lu, X. X., Kumm, M. & Oeurng, C. Reappraisal of sediment dynamics in the Lower Mekong River, Cambodia. *Earth Surf. Process. Landf.* **39**, 1855–1865 (2014).
  25. Wang, J. J., Lu, X. X. & Kumm, M. Sediment load estimates and variations in the lower Mekong River. *River Res. Appl.* **27**, 33–46 (2011).
  26. Darby, S. E., Leyland, J., Kumm, M., Räsänen, T. A. & Lauri, H. Decoding the drivers of bank erosion on the Mekong river: the roles of the Asian monsoon, tropical storms, and snowmelt. *Wat. Resour. Res.* **49**, 2146–2163 (2013).
  27. Manh, N. V., Dung, N. V., Hung, N. N., Merz, B. & Apel, H. Large-scale suspended sediment transport and sediment deposition in the Mekong delta. *Hydrol. Earth Syst. Sci.* **18**, 3033–3053 (2014).
  28. Kontgis, C., Schneider, A. & Ozdogan, M. Mapping rice paddy extent and intensification in the Vietnamese Mekong River Delta with dense time stacks of Landsat data. *Remote Sens. Environ.* **169**, 255–269 (2015).
  29. Brunier, G., Anthony, E. J., Goichot, M., Provansal, M. & Dussouillez, P. Recent morphological changes in the Mekong and Bassac river channels, Mekong Delta: the marked impact of river-bed mining and implications for delta destabilisation. *Geomorphology* **224**, 177–191 (2014).
  30. Redmond, G., Hodges, K. I., Mcsweeney, C., Jones, R. & Hein, D. Projected changes in tropical cyclones over Vietnam and the South China Sea using a 25 km regional climate model perturbed physics ensemble. *Clim. Dyn.* **45**, 1983–2000 (2015).
  31. Nittrover, J. A. & Viparelli, E. Sand as a stable and sustainable resource for nourishing the Mississippi River delta. *Nat. Geosci.* **7**, 350–354 (2014).
- Acknowledgements** This study was supported by awards NE/J021970/1, NE/J021571/1 and NE/J021881/1 from the UK Natural Environmental Research Council (NERC) and the Academy of Finland funded project SCART (grant number 267463). We thank the Mekong River Commission for access to hydrological and suspended sediment data and the Department for Hydrology and Water Resources in Cambodia for aDcp data and their logistical support. J.L.B. was also in receipt of a University of Southampton Diamond Jubilee Fellowship and National Great Rivers Research and Education Centre Fellowship that aided completion of this work.
- Author Contributions** S.E.D., J.L., C.R.H., D.R.P., J.L.B., A.P.N. and R.A. jointly conceived the study. C.R.H., S.E.D., J.L., J.L.B. and D.R.P. collected and processed the field data. C.R.H. constructed the sediment rating curves and, with S.E.D., undertook the data analysis. M.K. and H.L. conducted the model simulations, with the TC track data and rainfall anomalies being computed by J.L. S.E.D. drafted the paper, which was then edited by all co-authors.
- Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.E.D. (S.E.Darby@soton.ac.uk).
- Reviewer Information** *Nature* thanks L. Giosan and the other anonymous reviewer(s) for their contribution to the peer review of this work.



## METHODS

**Hydrological model.** The VMod hydrological model<sup>32</sup> was selected on the basis of its success in previous studies of the Mekong River basin<sup>26,33,34</sup>. As implemented for the Mekong River, VMod employs a  $5\text{ km} \times 5\text{ km}$  ( $25\text{ km}^2$ ) grid, with surface elevation, gradient, aspect, vegetation and soil type in each cell being extracted from the SRTM DEM<sup>35</sup>, GLC2000 land cover<sup>36</sup> and FAO soil type<sup>37</sup> data sets, respectively.

VMod simulations were forced using daily rainfall and temperature data estimated from a network of 151 meteorological stations (Fig. 1a). Specifically, the precipitation data employed here are from the Mekong River Commission (MRC) hydrometeorological database<sup>38</sup>, supplemented with Global Surface Summary of the Day (GSOD) data<sup>39</sup> for the Chinese part of the basin. These data have been carefully quality controlled<sup>33</sup>, and the MRC data therefore represent the highest-quality available data, with the best density of precipitation stations. However, as is frequently the case in developing nations, resource constraints have meant that there has not yet been a more recent release of the MRC product, constraining our study to the period 1981–2005. However, also pertinent to this choice of study period is the fact that in 2005 the total active storage of all dams on the Mekong was  $7.2\text{ km}^3$ , of which the active storage of Chinese dams was only  $0.8\text{ km}^3$ , meaning that the potential impact of dams is still rather minor at this date<sup>8</sup>. In contrast, by the year 2015 these figures had increased to  $\sim 55\text{ km}^3$  and  $24\text{ km}^3$ , respectively.

Estimates of daily rainfall totals and temperatures within each VMod grid cell were obtained by interpolating from the three nearest observations using inverse distance squared weighting. For daily rainfall totals, a multiplicative elevation correction (with coefficient  $0.0002\text{ mm m}^{-1}$ ) was employed to account for differences of elevation between each observation point and the location of the grid cell, whereas the temperature data were corrected for elevation using a lapse rate of  $-0.006\text{ K m}^{-1}$ . VMod simulates snowmelt using a degree-day model, in which the amount of snowmelt is obtained from daily average temperature exceeding a given threshold multiplied by a snowmelt coefficient  $K_{\text{melt}}$ . The model also computes snow evaporation, snowpack water storage, and refreezing. The snowmelt parameters employed here were calibrated in a previous study<sup>33</sup> using flow measurements at the Chiang Saeng gauging station. Glacier melt is computed similarly to snowmelt, albeit using a different set of parameters and the assumption of infinite storage.

In VMod the flow discharge is routed along a river network that is generated using DEM and map data. Each model grid cell has a river, either starting at that grid cell or flowing through it, to which the runoff from the cell is added. Flow within the river network is computed using a one-dimensional river model with a kinematic wave approximation. In this way simulated runoff at any point in the network reflects both the local and upstream contributions of precipitation, with the precipitation being deconstructed into cyclone and non-cyclone components as described later. In this approach the flow discharge (and hence sediment transport) that is attributable to cyclone and non-cyclone rainfall components at a given location in the river network is not explicitly parsed out as being attributable to a specific rainfall event. Instead, the simulated runoff components reflect the integrated effects of series of rainfall events that are delivered over longer time periods. Note that in the flow routing process, river cross-sections are represented using two superimposed trapezoids, with the lower one representing the main channel and the upper the floodplain, allowing for a representation of the effects of overbank storage on downstream attenuation of the flood wave.

Figure 2 (for Kratie, along with the left-hand panels of Extended Data Fig. 3 for the other hydrological stations) shows a comparison of simulated VMod versus observed runoff regimes at each of the gauging stations employed in this study. Note that, for clarity, Fig. 2 shows data only for the period 1995–2000, a period that includes the years that are most and least affected by TCs, but the goodness-of-fit measures reported here are for the entire simulation period (1 May 1981 to 31 March 2005). The four goodness-of-fit measures used are: (1) the mean discrepancy ratio for daily flows ( $Me$ ), which is the average of all the ratios (computed at each daily time step) of simulated to observed daily water flows, with  $Me = 1$  indicating perfect agreement between simulated and observed data; (2) the mean discrepancy ratio for annual peak flows ( $Me_p$ ); (3) the root mean square error (r.m.s.e.); and (4) the Nash–Sutcliffe Index (NSI)<sup>40</sup>. On the basis of these metrics (Extended Data Fig. 3), VMod, on average, under-predicts daily water flows throughout the study reach, while under-predicting the annual flood maxima in the lower parts (Stung Treng and Kratie) and over-predicting annual flood maxima in the upper parts (Luang Prabang, Mukdahan and Pakse) of the reach (Fig. 2 and Extended Data Fig. 3). Nevertheless, with NSI values varying between 0.749 (Luang Prabang) and 0.922 (Pakse), the overall performance of VMod is either ‘very good’ (Luang Prabang, Mukdahan, Stung Treng) or ‘excellent’ (Pakse, Kratie), on the basis of the classification scheme of Henriksen *et al.*<sup>41</sup>

**Rainfall scenarios and TC climatology.** The hydrological model as described earlier was run with two rainfall scenarios. The first ‘baseline’ scenario replicated actual conditions in the 1981–2005 study period and employed observed rainfall

totals. In the second scenario, these baseline totals were revised downwards by removing the rainfall estimated to have been delivered by tropical cyclones. The simulated runoff associated with tropical cyclones ( $Q_{\text{sim,TC}}$ ) was then computed by differencing the daily flows simulated under the two scenarios.

To estimate rainfall totals associated with TCs, we first employed the IBTrACS (version v03r02) storm tracks database<sup>42</sup> to locate the paths, at daily time steps, of all recorded TCs intersecting or passing near the Mekong Basin during 1981–2005. Rainfall anomalies associated with these storm paths were then defined by first interpolating, using the nearest neighbour, daily rainfall values observed at the network of 151 stations used in the baseline rainfall scenario onto a  $0.1^\circ$  ( $\sim 11\text{ km}^2$ ) resolution grid. Next, all rainfall stations located within a 500 km Haversine search radius<sup>43,44</sup> from the centroid of the storm on that date were identified. These identified stations were then temporarily (for the specific time step) removed from the analysis and an updated rainfall surface (minus the identified stations) was re-interpolated onto the same  $0.1^\circ$  grid. A rainfall anomaly surface, representing estimated rainfall associated with the identified storm and time step, was obtained by differencing the original and updated surfaces. This process was repeated for each daily time step, allowing the observed rainfall series at each meteorological station to be adjusted by subtracting rainfall anomalies within the grid square specific to each gauge from the observed daily rainfall totals. Note that since the hydrometeorological database we used in this analysis does not discriminate between precipitation associated or not associated with TCs, it is not possible to validate our estimates of cyclone-derived precipitation. For this reason, our estimates of rainfall associated with TCs are deliberately based on a method (nearest neighbour interpolation) that is more conservative than previous studies<sup>44</sup>, which simply assume that all rainfall within the assigned search radius is related to tropical cyclones. By the same token, while acknowledging that there is uncertainty regarding the typical radii of tropical cyclones, our decision to employ a 500 km search radius is again conservative in that it is at the lower end of the range of values typically used in previous studies<sup>45</sup>.

The IBTrACS data on which the above analyses are founded comprise six hourly best-track positions and intensity estimates. Only storms designated as in a tropical phase with one-minute maximum sustained surface wind speeds exceeding 34 knots ( $17.5\text{ m s}^{-1}$ ) are included in our analysis. The IBTrACS data were also used to compute the accumulated cyclone energy (ACE) metric<sup>46</sup> that we employ to characterize the TC climatology over the Mekong River basin for the period 1981–2005. The ACE parameter is analogous to the power dissipation index (PDI)<sup>47</sup> in that it convolves intensity and duration information for each individual TC observed in a defined area (here the sub-basins for the five gauging stations that are the focus of this study), offering considerable advantages over definitions based on the more familiar categorizations on the basis of wind speed<sup>48</sup>. In this context, our estimates of ACE are obtained by squaring the 6-hourly intensity estimates reported in the best-track database and integrating over the 1981–2005 study period.

**Sediment rating curves.** Sediment rating curves of the form:

$$C = aQ^b \quad (2)$$

were constructed for each hydrological station on the Mekong River mainstem below the China–Laos border and upstream of the Mekong Delta by fitting observed suspended solids concentration (SSC;  $C$ ) and observed water discharge ( $Q$ ) data (Extended Data Table 3) using nonlinear estimation techniques constructed using the Curve Fitting Toolbox in Matlab version R2014a. Specifically, a nonlinear least-squares power-law solver with one term was applied to the raw data, using the Trust-Region algorithm. The use of the power-law solver follows previous work<sup>49–51</sup> in optimizing the fit at the higher values of discharge and concentration that dominate overall transport. This procedure results in a poor fit for low discharges at Pakse (Extended Data Fig. 2) but using an alternative solver, designed to improve the low fit, is not justified. This is because doing so makes only a very minor ( $<2\%$ ) difference in the mean annual sediment load at Pakse while introducing substantial errors into the more important high-flow fits at the other stations. Note that our focus on suspended, rather than total, sediment load is not problematic since bed load is less than 20% of the total load (on the basis of comparisons of rivers from the data compilation of Turowski *et al.*<sup>52</sup> with suspended sediment concentrations similar to those of the Mekong River).

In terms of the data sources feeding into the sediment rating curves (Extended Data Table 3), at Luang Prabang, Mukdahan and Pakse the SSC and water discharge data were obtained from hydrological records archived by the MRC (available to download from <http://portal.mrcmekong.org/index>). However, the MRC SSC measurements are available only sporadically and have been acquired using a range of methodologies (reflecting the different approaches taken by differing hydrological agencies in this trans-national river) at the different gauging stations (Extended Data Table 3). All of the MRC’s SSC measurements

at Mukdahan were collected using United States Geological Survey (USGS)-designed isokinetic depth-integrated samplers (USGS D49 samplers) deployed at three verticals over the cross-section. The three samples are composited to provide a single sample from which the SSC is determined<sup>50</sup>. For the stations in Laos (that is, Luang Prabang and Pakse), the MRC SSC data were initially (1961) collected for a brief period using the same procedures as at Mukdahan, but subsequently the depth-integrated samplers were replaced with USGS P61 point-integrating samplers. To avoid potential problems with mixed sampling protocols in the data sets, and because depth-integrated sampling relies heavily on the even ascent of the sampler through the water column, to avoid biasing the SSC we excluded the relatively few data obtained using depth-integrated samplers from further consideration. The point-integrated samplers were deployed at three verticals over the cross-section, at heights of 0.2, 0.5 and 0.8 of the flow depth in the case of the point-sample (producing nine individual samples, from which the mean SSC for the cross-section is obtained by simple averaging). However, as shown in Extended Data Fig. 7, because the concentration of suspended sediment varies, both through the water column and laterally over the cross-section, simple averaging of point-based samples systematically biases the resulting estimate of the cross-section averaged SSC (relative to that obtained from alternative quasi-synoptic sampling techniques). We corrected for this effect by reducing the SSC values recorded within the MRC database by 26% for all the Laos and Thai stations (Extended Data Fig. 7). We derived this correction factor by comparing the averaged cross-section SSC computed from acoustic Doppler current profiler (aDcp) surveys in Cambodia, these aDcp surveys being undertaken as part of an aDcp field calibration exercise designed to retrieve SSC data from aDcp records archived by the Cambodian hydrological agency.

For the stations at Stung Treng and Kratie, sediment rating curves were constructed using flow discharge and SSC data (Extended Data Table 3) retrieved from the archives of the Cambodian Department of Hydrology and Water Resources (DHRW). These DHRW data were acquired via deployments of a four-beam 600 kHz aDcp (RD Instruments) during routine surveys undertaken in the period 2009 to 2014 by DHRW personnel. These aDcp surveys do not directly record suspended solids concentrations, but rather the archived DHRW data files contain acoustic backscatter (ABS) information recorded during the original surveys. We retrieved SSCs from these ABS data by means of a calibration function (Extended Data Fig. 7) that we derived on the basis of 54-point measurements of SSC deployed contemporaneously with the DHRW aDcp to record coeval ABS values in the same parcel of water following past guidelines<sup>53–55</sup>. In this field calibration procedure, the SSC data were obtained by filtering (Whatman GF/C glass microfibre grade 47 mm diameter 1.2 µm filter paper) and weighing the mass of solids retained from water samples collected at a wide range of flow depths and channel locations using a 31 Van Dorn sampler<sup>56</sup> during fieldwork that was spread over a wide range of flow conditions during 2013 and 2014. Consequently, the calibration function encompasses a wide range of SSC and ABS data. Analysis of ABS values and the suspended sediment grain size collected from the point samples reveals there is no relationship between the two, probably owing to the narrow range of grain sizes within the LMR<sup>57</sup>. Since the aDcp data provide a quasi-synoptic (less a blanking zone of 0.5 m at the top of the water column and a side-lobe interference zone of 10% of the flow depth at the bottom of the water column) image of ABS over the channel cross-section, the calibration function can be used to transform the ABS data to an accurate estimate of section-averaged SSC (Extended Data Fig. 7), as also noted earlier.

Having derived the rating curves for each gauging station (Extended Data Fig. 2), we then explicitly investigated whether the rating curves exhibit hysteresis effects associated with sediment exhaustion, which might be expected to lead to lower SSC values for a given discharge on the falling versus rising stages of the annual flood wave. However, no such evidence of hysteresis was identified (see Extended Data Fig. 2), presumably as a result of fluctuations in SSC being subdued owing to the large catchment areas and consequent effects of channel and floodplain storage in attenuating the peaks<sup>25</sup>.

We also considered whether there is a shift in sediment transport during flows affected by TCs, for example as a result of increased sediment supply from catchment erosion during storms. Specifically, we evaluated whether there are differences in sediment rating curves for flows that are (using the VMod model outputs to identify TC-affected flows and then cross-matching to identify SSC measurements that are TC affected) or are not affected by TCs. As indicated in Extended Data Table 3, this enabled us to identify 34 SSC samples during TC affected flows at Luang Prabang (14% of all observations at that station), while 30 samples were identified during TC affected flows at Mukdahan (3% of observations). We found that there were no significant (analysis of variance (ANOVA),  $P > 0.05$ ) differences between sediment ratings developed using the TC-affected versus the non-TC affected SSC data at either station. This indicates that we can

with confidence apply single rating curves for these stations, for both TC-affected and TC-unaffected flows. Since we are only able to discriminate TC-affected flows from VMod outputs during the 1981–2005 study period, and because there are no SSC data from this period at Stung Treng and Kratie, and there are too few SSC data at Pakse to identify any TC-affected measurements, there are no data to complete a similar formal analysis at these other three stations (Extended Data Table 3). Nevertheless, the very tight fit of these three stations' ratings (Extended Data Fig. 2), alongside the point that these stations are TC-affected during the period of SSC data collection, indicates that any shift in sediment transport processes during TCs is unlikely to have any material effects on the estimation of suspended solids loads at these locations.

Bearing in mind the relatively long periods over which the SSC data used to construct the sediment ratings at Luang Prabang, Mukdahan and Pakse were collected (Extended Data Table 3), we also tested for the possibility that varying ENSO phase, a known cause of hydroclimatological variability in the Mekong River, may lead to non-stationarity in the SSC values at these stations<sup>58,59</sup>, using dummy-variable regression analysis. Letting  $Z = 1$  if ENSO phase is positive (that is, El Niño) and 0 otherwise, then for the slope of the regression:

$$y = \beta_0 + \beta_1 ZX + \beta_2 X + \varepsilon \quad (3)$$

$$y = \begin{cases} \beta_0 + (\beta_1 + \beta_2)X + \varepsilon & \text{if ENSO phase is positive} \\ \beta_0 + \beta_2 X + \varepsilon & \text{if ENSO phase is negative} \end{cases} \quad (4)$$

Then, for the intercept of the regression:

$$y = \beta_0 + \beta_1 Z + \beta_2 X + \varepsilon \quad (5)$$

$$y = \begin{cases} (\beta_0 + \beta_1) + X + \varepsilon & \text{if ENSO phase is positive} \\ \beta_0 + X + \varepsilon & \text{if ENSO phase is negative} \end{cases} \quad (6)$$

Or, for both slope and intercept:

$$y = \beta_0 + \beta_1 Z + \beta_2 X + \beta_3 + \varepsilon \quad (7)$$

$$y = \begin{cases} (\beta_0 + \beta_1) + (\beta_2 + \beta_3)X + \varepsilon & \text{if ENSO phase is positive} \\ \beta_0 + \beta_3 X + \varepsilon & \text{if ENSO phase is negative} \end{cases} \quad (8)$$

We found no significant difference at the 0.05 significance level (ANOVA on dummy-variable regression coefficients for each site) in the SSCs, for a given  $Q$ , as a function of ENSO phase, demonstrating that there is therefore no evident bias in the SSCs introduced as a function of climate variability associated with ENSO. With the completion of the first major main-stem cascade of dams on the Chinese portion of the Mekong River in 1993, we also considered whether the SSC data differ pre- and post-1993. Accordingly, a similar analysis (equations (3) to (8)) was conducted for those sites (Luang Prabang and Mukdahan) at which SSC samples span the pre- and post-dam periods. We found that at Mukdahan no significant difference exists at the 0.05 significance level (ANOVA on dummy-variable regression coefficients), implying that there is no reason to split the data based on the pre- and post-dam periods. However, a significant difference ( $P < 0.05$ ) between the pre- and post-dam periods does exist at Luang Prabang (ANOVA test statistic = 9.7377,  $n = 236$ , degrees of freedom (df) = 1,232). Consequently, at Luang Prabang, we calculate suspended solids loads (see later) using the pre- and post-dam rating curves (Extended Data Fig. 2) for the periods 1981–1992 and 1993–2005, respectively. Finally, we emphasize that our analysis does not account for anthropogenic factors, such as flow regulation through reservoirs, land-use or land cover change, or increasing sediment mining, which could potentially introduce a trend into the relationships between flow discharge and SSC at each gauging station. Our suspended sediment rating curves therefore assume stationarity of these factors over the 1981–2005 study period.

**Sediment load estimation.** The lack of hysteresis and apparent stationarity of the SSC data means that we were able to employ a single (two at Luang Prabang, one for the pre- and one for the post-dam periods) sediment rating curve specific to each station (Extended Data Fig. 2), together with the continuous water discharge records obtained from our hydrological modelling, to estimate daily suspended solids loads (Fig. 2 and Extended Data Fig. 3) for the 1981–2005 study period. These daily loads were in turn used to compute, by summation, the annual sediment loads for each station (Fig. 3 and Extended Data Fig. 4). Note that since the modelling period extended from 1 May 1981 to 31 March 2005, we report annual sediment loads only for those years (1982 to 2004 inclusive) for which full-year records are available. Mean annual suspended solids loads for each station over the 22-year

period (1982 to 2004) were then obtained by calculating the arithmetic mean of these annual loads (Extended Data Table 2).

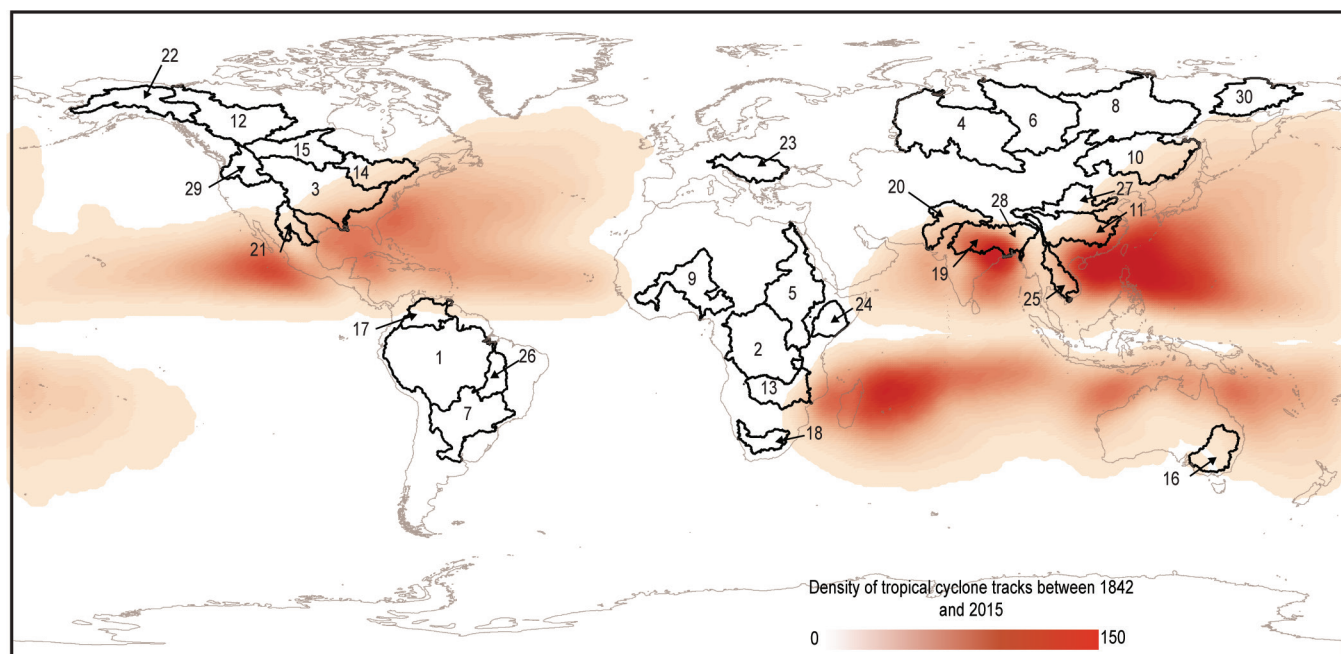
**Statistical analysis.** No statistical methods were used to predetermine sample size. Mann–Kendall<sup>60</sup> tests, used to evaluate whether there are significant (at 95% confidence) temporal trends (the magnitude of the trend being equated to Sen's slope, with uncertainty equated to the 95% confidence bounds on the Sen slope estimates) in the computed annual sediment loads, were computed in Matlab R2014a using the ktaub.m file written by J. Burkey (2006), which is available from the Matlab Exchange (<http://www.mathworks.com/matlabcentral/fileexchange/11190-mann-kendall-tau-b-with-sen-s-method-enhanced-/content/ktaub.m>).

**Data.** The precipitation and temperature data used in the hydrological model simulations are taken from the MRC hydrometeorological database<sup>38</sup> (not available online) supplemented with GSOD data<sup>39</sup> for the Chinese part of the basin (<ftp://ftp.ncdc.noaa.gov/pub/data/gsod/>; years 1981–2005). The IBTrACS (version v03r02) storm tracks database<sup>42</sup> that we used to estimate the track locations and hence precipitation anomalies associated with TCs was downloaded from the IBTrACS website (<https://www.ncdc.noaa.gov/ibtracs/index.php?name=ibtracs-data>; IBTrACS-All data v03r02 all storms line shapefile). Note that we are not able to make the input data files used in the hydrological model simulations available as the precipitation and temperature data are from the MRC (as described earlier) under a licence that precludes redistribution of products or derived products. Water discharge data used in the validation of the hydrological model are from the hydrological records archived in the MRC data portal (<http://portal.mrcmekong.org/index> as discharge records from Luang Prabang (station identifier 011201; unique data set accession 21301), Mukdahan (station identifier 013402; unique data set accession 3301), Pakse (station identifier 013901; unique data set accession 3141), Stung Treng (station identifier 014501; unique data set accession 2809), and Kratie (station identifier 014901; unique data set accession 2811)), as are the suspended sediment concentration data (available from <http://portal.mrcmekong.org/index> as sediment concentration records from: station identifier 011201, unique data set accession 4746; station identifier 013402, unique data set accession 4849; and station identifier 013901, unique data set accession 4773, respectively) used to derive the sediment rating curves at Luang Prabang, Mukdahan and Pakse. The aDcp data files used to derive the sediment rating curves for the stations at Stung Treng and Kratie are available on request from the Cambodian Department of Hydrology and River Works (DHRW; <http://www.dhrw-cam.org/index.php>).

**Code sharing.** The VMod hydrological model software as employed in this study is available to download from <http://www.eia.fi/vmod>. The related analytical code comprises the bespoke Matlab scripts, authored by J.L., that were used to partition out the cyclone-influenced rainfall. These scripts are not publicly available as they are currently being developed and used in commercial applications.

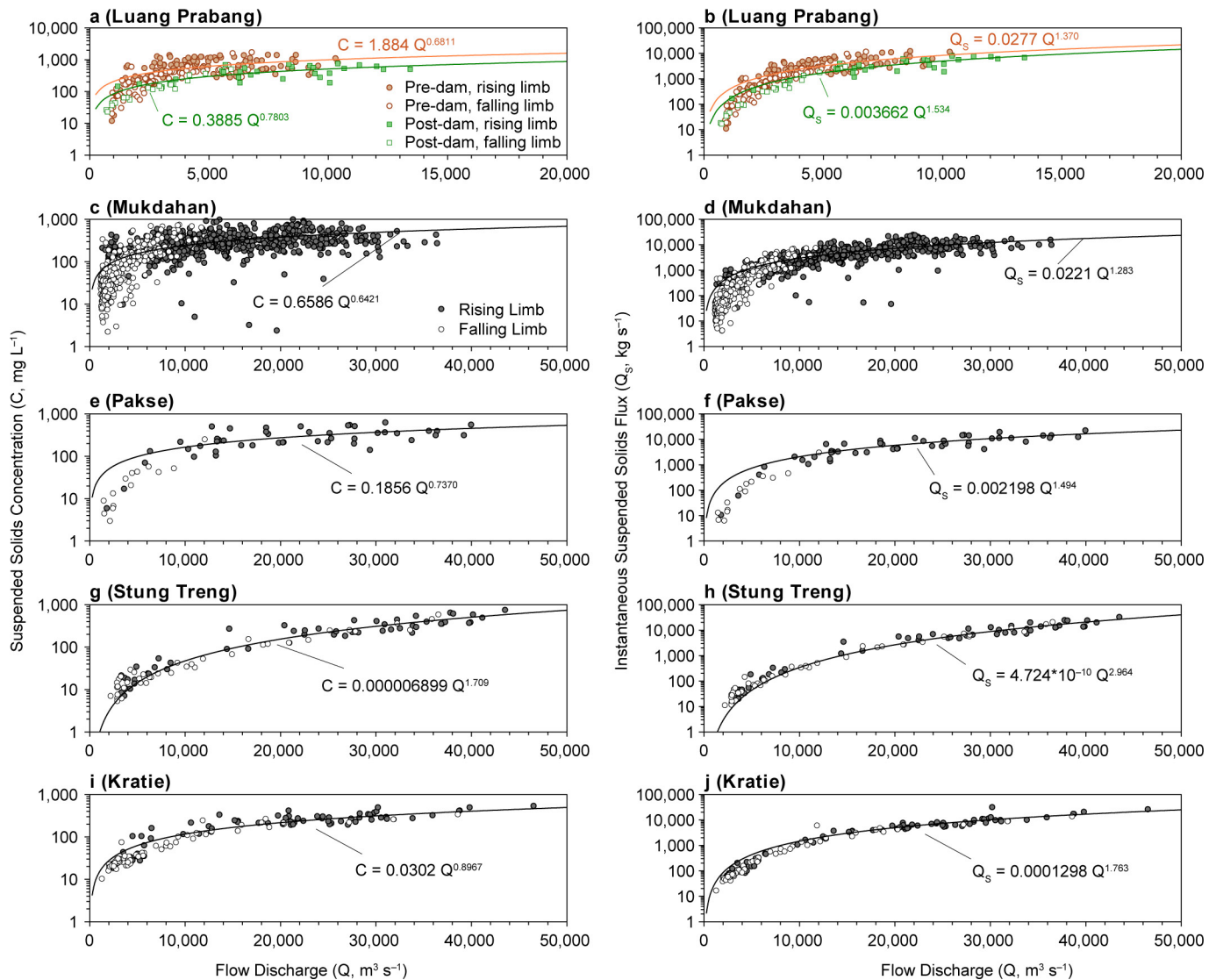
32. Koponen, J. H. *et al.* *HBV and IWRM Watershed Modelling User Guide* (MRC Information and Knowledge Management Programme, 2010).
33. Lauri, H. *et al.* Future changes in Mekong River hydrology: impact of climate change and reservoir operation on discharge. *Hydrol. Earth Syst. Sci.* **16**, 4603–4619 (2012).
34. Lauri, H. *VMod 5km Grid Hydrological Modeling Report* (EIA Ltd.) (Aalto Univ., 2009).
35. Jarvis, A. H. *et al.* *Hole-Filled Seamless SRTM Data Version 4* (The CGIAR Consortium for Spatial Information, 2008).
36. IES. *Global Land Cover 2000* (IES, 2000).
37. Food and Agricultural Organization of the United Nations. *WRB Map of World Soil Resources* (FAO, 2003).
38. Mekong River Commission. *Hydrometeorological Database of the Mekong River Commission* (Mekong River Commission, 2011).
39. US National Climatic Data Center. *Global Surface Summary of the Day (GSOD)* (NCDC, 2010).
40. Nash, J. E. & Sutcliffe, J. V. River flow forecasting through conceptual models part I—a discussion of principles. *J. Hydrol. (Amst.)* **10**, 282–290 (1970).
41. Henriksen, H. J. *et al.* Assessment of exploitable groundwater resources of Denmark by use of ensemble resource indicators and a numerical groundwater–surface water model. *J. Hydrol. (Amst.)* **348**, 224–240 (2008).
42. Knapp, K. R. *et al.* The international best track archive for climate stewardship (IBTrACS): Unifying tropical cyclone best track data. *Bull. Am. Meteorol. Soc.* **91**, 363–376 (2010).
43. Rodgers, E. B. *et al.* Contribution of tropical cyclones to the North Pacific climatological rainfall as observed from satellites. *J. Appl. Meteorol.* **39**, 1658–1678 (2000).
44. Englehart, P. J. & Douglas, A. V. The role of eastern North Pacific tropical storms in the rainfall climatology of western Mexico. *Int. J. Climatol.* **21**, 1357–1370 (2001).
45. Kubota, H. & Wang, B. How much do tropical cyclones affect seasonal and inter-annual rainfall variability over the Western North Pacific? *J. Clim.* **22**, 5495–5510 (2009).
46. Bell, G. D. *et al.* Climate assessment for 1999. *Bull. Am. Meteorol. Soc.* **81**, s1–s50 (2000).
47. Emanuel, K. Increasing destructiveness of tropical cyclones over the past 30 years. *Nature* **436**, 686–688 (2005).
48. Webster, P. J., Holland, G. J., Curry, J. A. & Chang, H. R. Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science* **309**, 1844–1846 (2005).
49. Ferguson, R. I. River loads underestimated by rating curves. *Wat. Resour. Res.* **22**, 74–76 (1986).
50. Walling, D. E. *Evaluation and Analysis of Sediment Data from the Lower Mekong River* (Mekong River Commission, 2005).
51. Walling, D. E. The changing sediment load of the Mekong River. *Ambio* **37**, 150–157 (2008).
52. Turowski, J. M., Rickenmann, D. & Dadson, S. J. The partitioning of the total sediment load of a river into suspended load and bedload: a review of empirical data. *Sedimentology* **57**, 1126–1146 (2010).
53. Kostaschuk, R. J. *et al.* Measurement of flow velocity and sediment transport with an acoustic Doppler current profiler. *Geomorphology* **68**, 25–37 (2005).
54. Szupiany, R. N. *et al.* Morphology, flow structure and suspended bed sediment transport at two large braid-bar confluences. *Wat. Resour. Res.* **45**, W05415 (2009).
55. Shugar, D. *et al.* On the relationship between flow and suspended sediment transport over the crest of a sand dune, Rio Parana, Argentina. *Sedimentology* **57**, 252–272 (2010).
56. Van Dorn, W. G. Large-volume water samplers. *Eos Trans. AGU* **37**, 682–684 (1956).
57. Bravard, J.-P., Goichot, M. & Tronchère, H. An assessment of sediment-transport processes in the Lower Mekong River based on deposit grain sizes, the CM technique and flow-energy data. *Geomorphology* **207**, 174–189 (2014).
58. Räsänen, T. & Kumm, M. Spatiotemporal influences of ENSO on precipitation and flood pulse in the Mekong River Basin. *J. Hydrol. (Amst.)* **476**, 154–168 (2013).
59. Ward, P. J. *et al.* Annual flood sensitivity to El Niño Southern Oscillation at the global scale. *Hydrol. Earth Syst. Sci.* **18**, 47–66 (2014).
60. Kendall, M. G. A new measure of rank correlation. *Biometrika* **30**, 81–93 (1938).





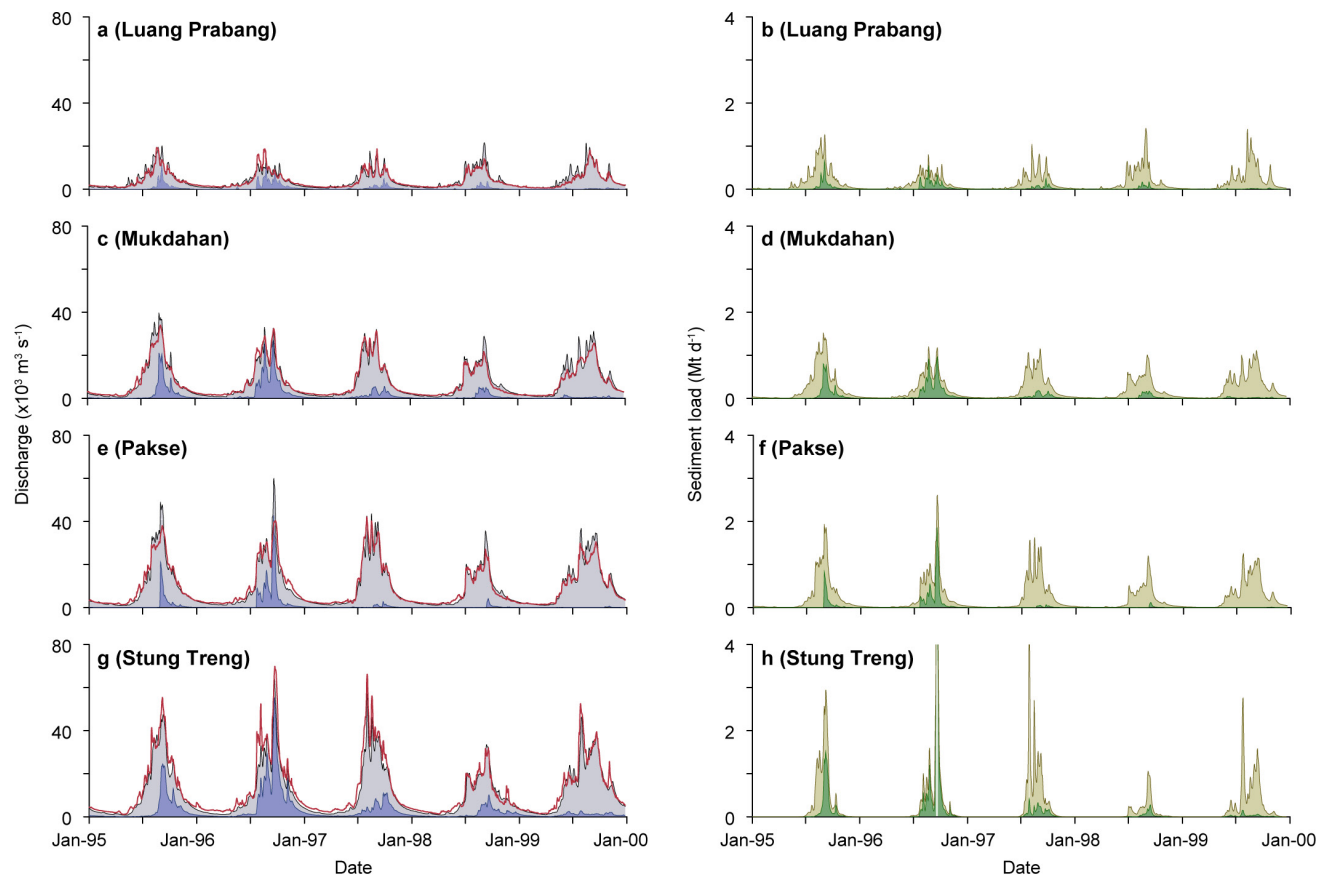
**Extended Data Figure 1 | Locations of the world's 30 largest (by drainage area) rivers.** The numbers identify the basins listed in Extended Data Table 1. Note that the Ganges (basin 19) and Brahmaputra (basin 28) catchments are outlined as a single basin in the figure. Also shown is the

density of all TC tracks from 1842 to 2015 as recorded in the IBTrACS<sup>42</sup> database. Track density was calculated using the point density function in ArcGIS 10.1.



**Extended Data Figure 2 | Sediment rating curves for the five river gauging stations on the Lower Mekong River.** **a, c, e, g, i,** The relationship between flow discharge ( $Q$ ) and suspended solids concentration ( $C$ ) at: Luang Prabang (pre-dam:  $n = 187$ ,  $r^2 = 0.338$ ; post-dam:  $n = 49$ ,  $r^2 = 0.648$ ) (**a**); Mukdahan ( $n = 1,159$ ,  $r^2 = 0.497$ ) (**c**); Pakse ( $n = 60$ ,  $r^2 = 0.591$ ) (**e**); Stung Treng ( $n = 95$ ,  $r^2 = 0.870$ ) (**g**); and Kratie ( $n = 140$ ,  $r^2 = 0.850$ ) (**i**). **b, d, f, h, j,** These panels show how the relationships in **a, c, e, g, i** and **i** propagate through to give the relationship between flow discharge ( $Q$ ) and instantaneous sediment load ( $Q_s$ ) at the same stations: Luang Prabang (pre-dam:  $n = 187$ ,  $r^2 = 0.791$ ; post-dam:  $n = 49$ ,  $r^2 = 0.864$ ) (**b**); Mukdahan ( $n = 1,159$ ,  $r^2 = 0.693$ ) (**d**); Pakse ( $n = 60$ ,  $r^2 = 0.780$ ) (**f**); Stung Treng ( $n = 95$ ,  $r^2 = 0.900$ ) (**h**); and Kratie ( $n = 140$ ,  $r^2 = 0.931$ ) (**j**). All the fits shown are significant at  $P < 0.00001$ . Note that the scales for **a** and **b**

(Luang Prabang) differ from those for the other panels. We recognize that the fits for  $Q$  versus  $Q_s$  in **b, d, f, h** and **j** are stronger than the fits between  $Q$  and  $C$  because of the auto-correlation arising when transforming  $C$  to  $Q_s$  ( $Q_s = C \times Q/1,000$ ). For the stations at Mukdahan, Pakse, Stung Treng and Kratie, a single rating curve is employed (black lines), as there is no evidence of hysteresis between the rising (filled circles) and falling (open circles) limbs of the hydrograph (see Methods). At Luang Prabang, there is likewise no evidence of hysteresis between the rising (coloured filled symbols) and falling (coloured open symbols) limbs. However, two rating functions are employed at this station, one for the pre-dam (orange coloured lines) and post-dam (green coloured lines) periods (see Methods).

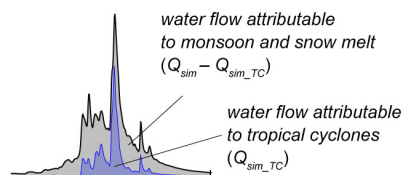


i

River Gauging Station	Goodness of Fit Measures			
	Mean Discrepancy Ratio – Daily Flows	Mean Discrepancy Ratio – Annual Peaks	Root Mean Square Error	Nash-Sutcliffe Index
	( $Me$ )	( $Me_p$ )	( $RMSE$ , $m^3 s^{-1}$ )	( $NSI$ )
Luang Prabang	0.99	1.19	1690	0.749
Mukdahan	0.87	1.08	3170	0.821
Pakse	0.89	1.15	2700	0.922
Stung Treng	0.73	0.84	6210	0.789
Kratie	0.77	0.89	5800	0.808

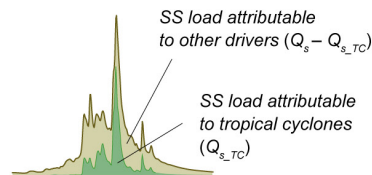
## DAILY WATER FLOW

- $Q_{sim}$  - simulated water flows
- $Q_{sim\_TC}$  - simulated water flows attributable to tropical cyclones
- $Q_{obs}$  - observed water flows



## DAILY SUSPENDED SEDIMENT LOAD

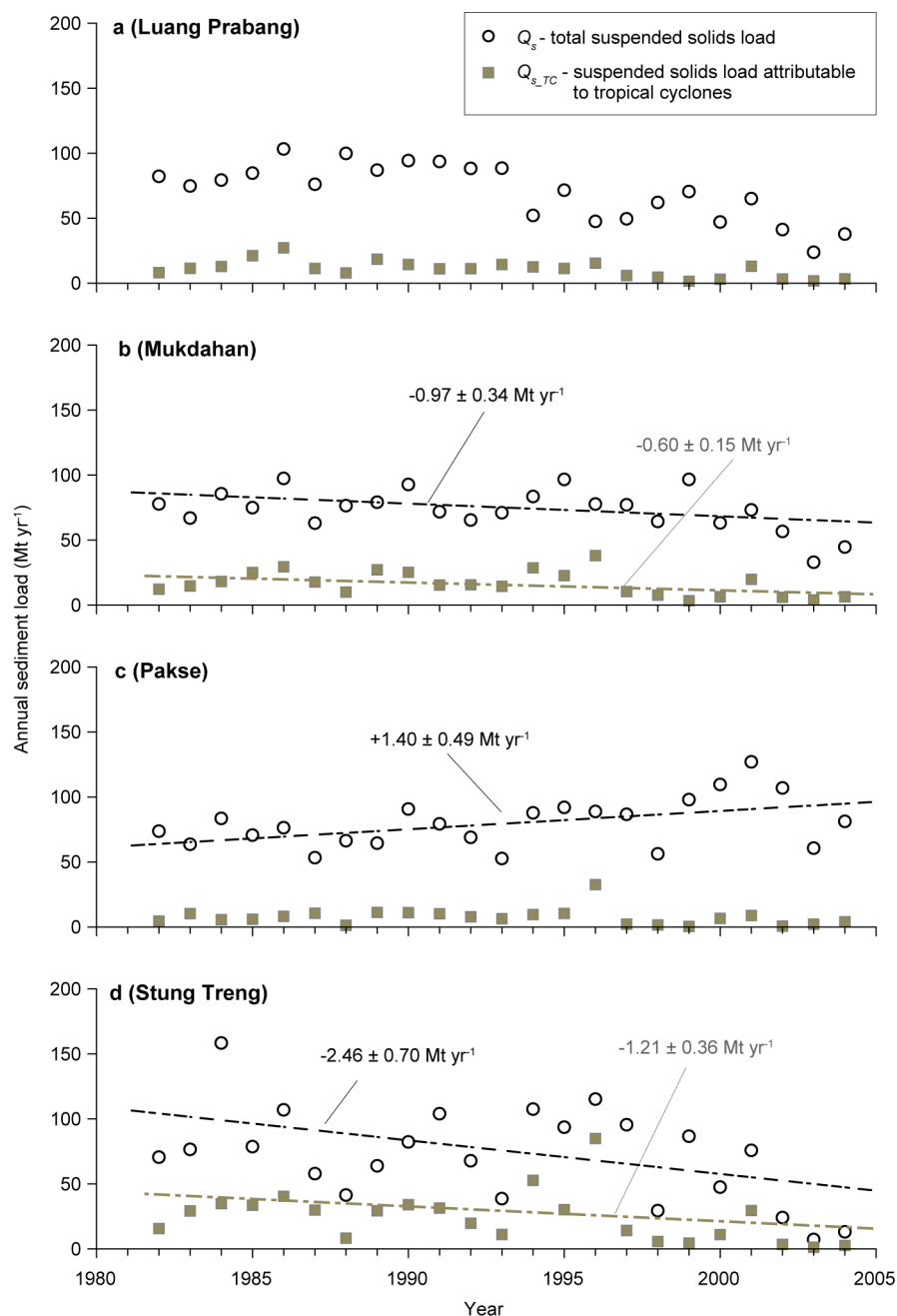
- $Q_s$  - total suspended solids load
- $Q_{s\_TC}$  - suspended solids load attributable to tropical cyclones



**Extended Data Figure 3 | Daily flow discharge and suspended solids load at selected Mekong River gauging stations from 1 January 1995 to 31 December 1999.** **a, c, e, g,** Daily simulated ( $Q_{sim}$ ) and observed ( $Q_{obs}$ ) water flows, along with the daily water flows attributable to tropical cyclones ( $Q_{sim\_TC}$ ) at Luang Prabang (**a**), Mukdahan (**c**), Pakse (**e**) and Stung Treng (**g**). **b, d, f, h,** Daily total suspended solids load ( $Q_s$ , in Mt per day) and daily suspended solids load attributable to TCs ( $Q_{s\_TC}$ ; also in Mt per day) at Luang Prabang (**b**), Mukdahan (**d**), Pakse (**f**) and Stung Treng (**h**). Note that the period 1995 to 1999 encompasses the years

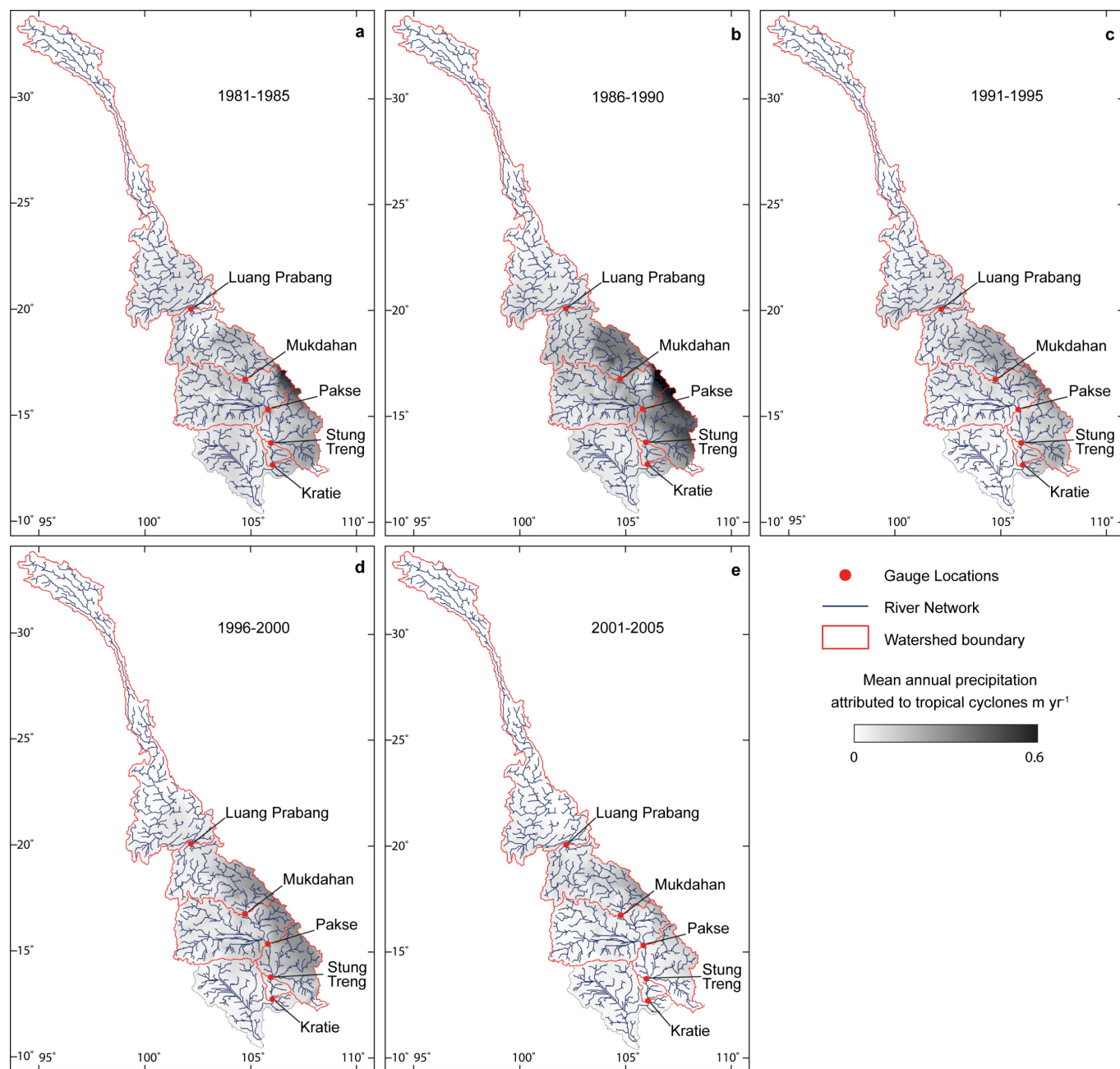
during the 1981–2005 study period that are the most (1996) and least (1999) strongly affected by TCs. **i,** Goodness-of-fit measures comparing VMod simulated and observed water flows at five river gauging stations on the Lower Mekong River. Note that the goodness-of-fit metrics are all based on the mean daily flows for the full simulation period (1 May 1981 to 31 March 2005), with the exception of the mean discrepancy ratio for the annual flood peaks ( $Me_p$ ). The  $Me_p$  metric is computed using the ratio of simulated maximum daily discharge to observed maximum daily discharge in each year of the record (1981–2004 inclusive) studied here.



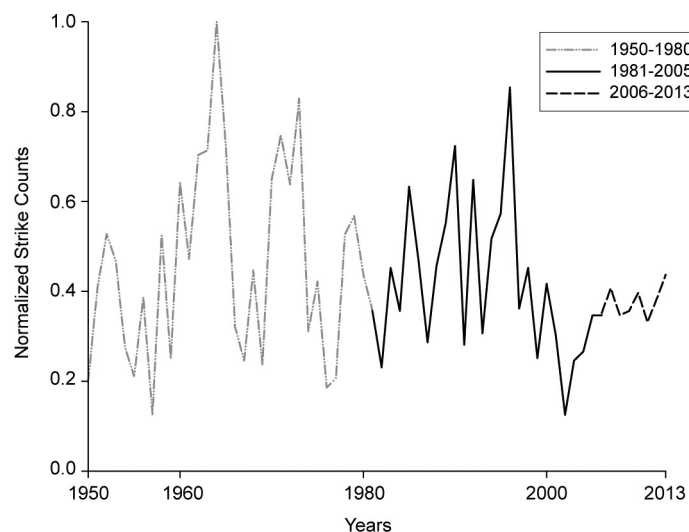


**Extended Data Figure 4 | Time series of annual suspended solids load at selected river gauging stations during 1982 to 2004.** a, Luang Prabang; b, Mukdahan; c, Pakse; d, Stung Treng. The symbols indicate the total suspended solids load ( $Q_s$ ; open circles) and suspended solids load

attributable to TCs ( $Q_{s-TC}$ ; filled squares). Significant ( $P \leq 0.05$ ) trends as identified by Mann–Kendall analysis are indicated by the dashed lines, with the corresponding time rate of change of annual suspended solids load annotated on the plot.



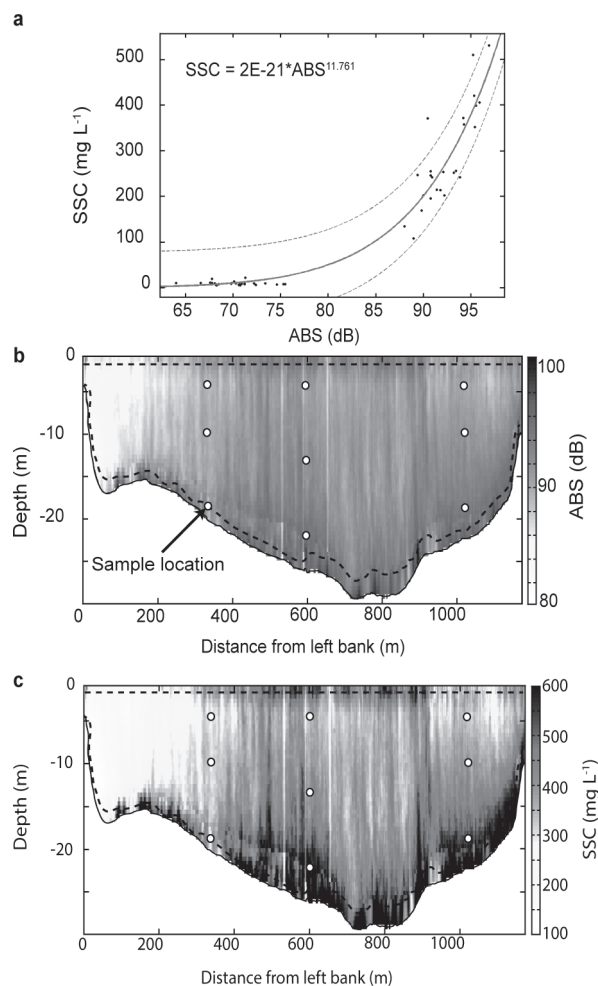
**Extended Data Figure 5 | Spatial distributions of mean annual rainfall contributed from TCs over the Mekong Basin. a, 1981–1985; b, 1986–1990; c, 1991–1995; d, 1996–2000; e, 2001–2005. Note the pronounced declines in rainfall associated with TCs at Stung Treng and Kratie in particular.**



**Extended Data Figure 6 | Strike counts for TCs tracking across the Mekong basin during 1950–2013.** The strike count data plotted are extracted from the IBTrACS<sup>42</sup> database and normalized by the maximum count (199) observed in 1964. We employ strike count, rather than precipitation, data in this longer-term historical analysis because reliable precipitation data are not available outside of the 1981–2005 period that is the main focus of the study. Similarly, mean wind speed data, which in principle could be used to estimate variations in ACE as a proxy for

precipitation, are available only sporadically outside of 1981–2005. In terms of strike counts, the data suggest that there is a periodicity in the long-term cyclone climatology, with the most recent data (2006–2013) having annual strike counts similar to the 1950–2013 mean of  $87 \pm 37$ . However, these data must be treated with caution since strike count data do not report the intensity or locations of cyclone tracks, both of which are important controls on the precipitation delivered to the basin by these TCs.





**Extended Data Figure 7 | Procedures used to determine cross-section mean suspended solids concentration from acoustic Doppler current profiler data.** **a**, Calibration function (solid line;  $n = 54$ ,  $r^2 = 0.9306$ ,  $P < 0.0001$ ) linking the suspended solids concentration (SSC) to acoustic backscatter (ABS) for the 600 kHz (RD Instruments) acoustic Doppler current profiler (aDcp) instrument employed in this study (dashed lines indicate 95% prediction intervals). **b**, Example of quasi-synoptic ABS field obtained from the aDcp survey at the Kratie gauging station on 23 September 2013 (flow discharge,  $Q = 57,000 \text{ m}^3 \text{ s}^{-1}$ ). Note that there is a small blanking distance close to the water surface and a zone of side-lobe interference near the bed (indicated by the dashed black lines) where no ABS values are returned, and the ABS values in these zones are therefore determined by interpolation. **c**, SSC field obtained on the basis of the ABS values in **b** and using the calibration function in **a**. Note how the locations of the nine point-based SSC estimates collected using the sampling procedure adopted at Luang Prabang and Pakse lead to a deviation of the cross-section mean SSC derived from the aDcp-estimated SSC field in **c** and the point-based sampling procedure. We compared 11 cross-section mean SSCs obtained using point-based versus aDcp sampling procedures at locations throughout the Mekong River south of Kratie to correct (by 26%) the consequent bias arising from cross-section averaging of point-based samples.

**Extended Data Table 1 | Characteristics of the world's 30 largest rivers with data from ref. 1**

Basin ID	Basin Name	Drainage Area (10 <sup>6</sup> km <sup>2</sup> )	Mean annual runoff (km <sup>3</sup> yr <sup>-1</sup> )	Mean annual sediment flux (Mt yr <sup>-1</sup> )	Accumulated Cyclone Energy (ACE, 10 <sup>4</sup> kn <sup>2</sup> )
1	Amazon	6.3	6300	1200	0.59
2	Congo	3.8	1300	43	0
3	Mississippi	3.3	490	120	10.06
4	Ob	3.0	390	16	0
5	Nile	2.9	30	0.2	0
6	Yenisei	2.6	620	4.1	0
7	Parana	2.6	530	90	1.75
8	Lena	2.5	520	20	0
9	Niger	2.2	160	40	0.12
10	Amur	1.9	350	52	1.12
11	Yangtze	1.8	900	470	12.14
12	Mackenzie	1.8	310	100	0
13	Zambezi	1.3	100	9	3.03
14	St. Lawrence	1.2	340	4.6	0
15	Nelson	1.1	89	n/a	0
16	Murray-Darling	1.1	7.9	1	0.65
17	Orinoco	1.1	1100	210	3.36
18	Orange	1.0	4.5	17	1.83
19	Ganges	0.98	490	520	4.50
20	Indus	0.98	5	10	1.66
21	Rio Grande	0.87	0.7	0.7	6.38
22	Yukon	0.85	210	54	0
23	Danube	0.82	210	42	0
24	Shebelle	0.81	19	n/a	0
25	Mekong	0.80	450	110	8.09
26	Tocantins	0.76	370	75	0
27	Yellow	0.75	15	150	0.91
28	Brahmaputra	0.67	630	540	2.46
29	Columbia	0.67	240	9.7	0
30	Kolyma	0.60	120	10	0

The identification numbers identify the locations of the drainage basins shown in Extended Data Fig. 1. The data indicate that the sediment loads from these 30 largest rivers (note that here we rank these large rivers on the basis of their drainage area) together sum to 3.92 billion tonnes per year, a substantial proportion (20.6%) of the total global riverine flux as estimated previously<sup>1</sup>. The tropical cyclone climatology for the period 1950–2013 is represented for each basin by the ACE<sup>46</sup> for each basin, with the underpinning data being extracted from the IBTrACS<sup>42</sup> database.

Extended Data Table 2 | Hydrometeorological data (1982–2004) for hydrological stations on the Lower Mekong River

River Gauging Station	Accumulated Cyclone Energy (ACE, $10^4 \text{ kJ m}^{-2}$ )	Rainfall (P, mm)	Rainfall due to Tropical Cyclones ( $P_{TC}$ , mm)	Runoff (Q, $\text{m}^3 \text{ yr}^{-1}$ )	Runoff due to Tropical Cyclones ( $Q_{TC}$ , $\text{m}^3 \text{ yr}^{-1}$ )	Suspended Solids Load ( $Q_s$ , Mt $\text{yr}^{-1}$ )	Suspended Solids Load due to Tropical Cyclones ( $Q_{s_{TC}}$ , Mt $\text{yr}^{-1}$ )	Proportion of Load Forced by Tropical Cyclones (%)
Luang Prabang	0.12	1157	21	3945	540	$70.4 \pm 21.8$	$10.7 \pm 6.4$	15.2
Mukdahan	2.84	1346	47	7230	1550	$73.4 \pm 15.9$	$16.5 \pm 9.3$	22.5
Pakse	3.49	1356	56	9155	770	$79.8 \pm 19.0$	$7.5 \pm 6.6$	9.3
Stung Treng	5.64	1436	67	10375	2960	$71.3 \pm 36.3$	$24.2 \pm 19.3$	34.0
Kratie	5.76	1438	67	10490	3020	$87.4 \pm 28.7$	$27.7 \pm 17.6$	31.7

Errors represent one standard deviation around the mean annual loads. The ACE for each station during the same period is also indicated.



**Extended Data Table 3 | Data sources used for sediment rating curves derived herein**

Gauging Station	Number of Samples	Period of Data Availability	Sampling Method
Luang Prabang	236 (34)	1986-1989; 1992; 1997-2000; 2002	Point-integrated
Mukdahan	1159 (30)	1962-1965; 1967-1980; 1982; 1991-1994; 1996-1997; 1999-2007	Depth-integrated
Pakse	60	1998-2002	Point-integrated
Stung Treng	95	2009-2014	aDcp backscatter
Kratie	140	2009-2014	aDcp backscatter

Number of samples refers to the total number of suspended solids concentration (SSC) data points used in the derivation of the sediment rating curves, with the numbers in parentheses indicating the number of SSC data points associated with TC-induced runoff events. The latter are defined herein as runoff events for which at least 25% of the runoff was associated with TC-induced runoff. Consequently, it is only possible to identify TC-affected SSC measurements in the 1981–2005 model simulation period. Note that no TC-induced runoff events were associated with the 60 SSC measurements made at Pakse during 1998–2002 and that the available data from Stung Treng and Kratie post-date the 1981–2005 study period.

# Cultural innovation and megafauna interaction in the early settlement of arid Australia

Giles Hamm<sup>1</sup>, Peter Mitchell<sup>2</sup>, Lee J. Arnold<sup>3</sup>, Gavin J. Prideaux<sup>4</sup>, Daniele Questiaux<sup>5</sup>, Nigel A. Spooner<sup>5,6</sup>, Vladimir A. Levchenko<sup>7</sup>, Elizabeth C. Foley<sup>1</sup>, Trevor H. Worthy<sup>4</sup>, Birgitta Stephenson<sup>8,9</sup>, Vincent Coulthard<sup>10</sup>, Clifford Coulthard<sup>10</sup>, Sophia Wilton<sup>10</sup> & Duncan Johnston<sup>10</sup>

**Elucidating the material culture of early people in arid Australia and the nature of their environmental interactions is essential for understanding the adaptability of populations and the potential causes of megafaunal extinctions 50–40 thousand years ago (ka). Humans colonized the continent by 50 ka<sup>1,2</sup>, but an apparent lack of cultural innovations compared to people in Europe and Africa<sup>3,4</sup> has been deemed a barrier to early settlement in the extensive arid zone<sup>2,3</sup>. Here we present evidence from Warraty rock shelter in the southern interior that shows that humans occupied arid Australia by around 49 ka, 10 thousand years (kyr) earlier than previously reported<sup>2</sup>. The site preserves the only reliably dated, stratified evidence of extinct Australian megafauna<sup>5,6</sup>, including the giant marsupial *Diprotodon optatum*, alongside artefacts more than 46 kyr old. We also report on the earliest-known use of ochre in Australia and Southeast Asia (at or before 49–46 ka), gypsum pigment (40–33 ka), bone tools (40–38 ka), hafted tools (38–35 ka), and backed artefacts (30–24 ka), each up to 10 kyr older than any other known occurrence<sup>7,8</sup>. Thus, our evidence shows that people not only settled in the arid interior within a few millennia of entering the continent<sup>9</sup>, but also developed key technologies much earlier than previously recorded for Australia and Southeast Asia<sup>8</sup>.**

Ten archaeological sites, between 41 and 28 kyr in age, have previously been recorded from arid Australia<sup>10</sup> (Fig. 1). Many lack well-stratified deposits and very few span durations of more than 20 kyr. The scarcity of late Pleistocene sites, especially in the southern arid interior, continues to prevent reliable interpretations of the nature, timing and implications of human colonization across the continent. Warraty rock shelter provides important new evidence for the early occupation of arid Australia. The site is an elevated rock shelter located in the Adnyamathanha country within the present arid zone at the northern end of the Flinders Ranges, in the southern Lake Eyre Basin, South Australia (Fig. 1). The shelter contains a stratified, 1-m-deep, intact archaeological deposit composed of four stratigraphic units (SU1–SU4; Fig. 2, Extended Data Figs 1–3 and Supplementary Information: stratigraphy).

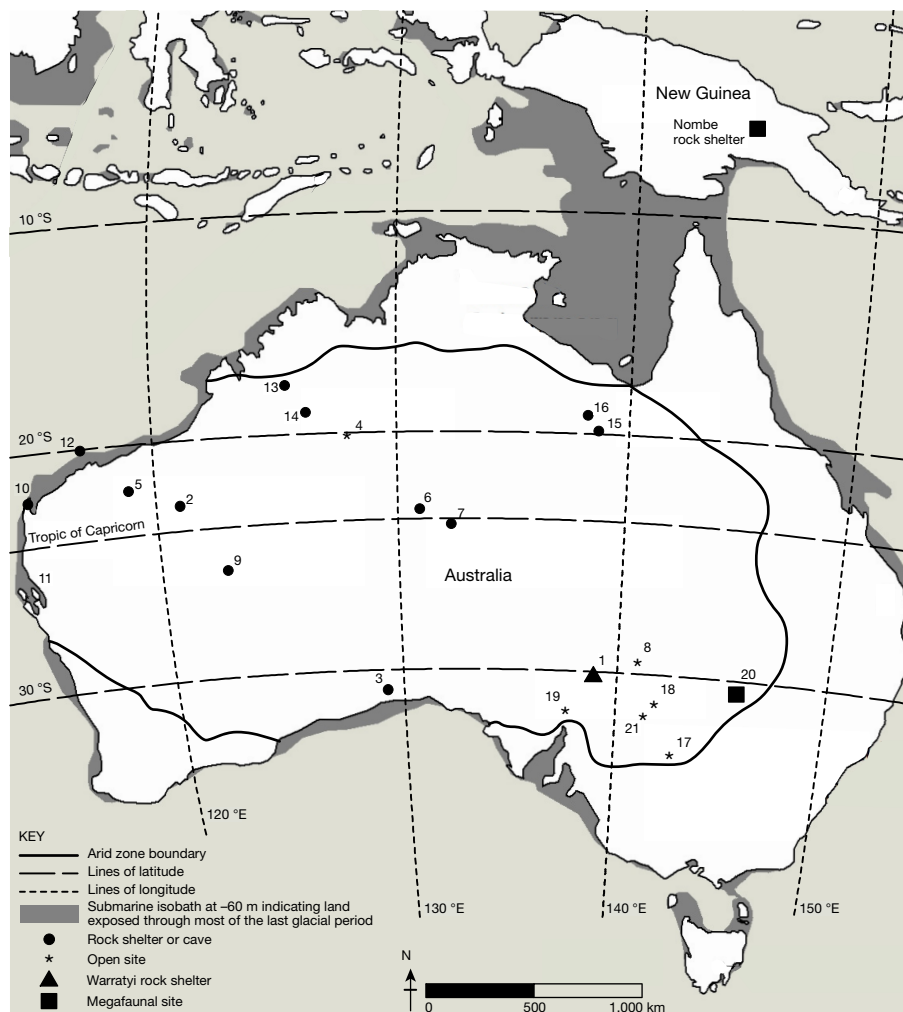
Single-grain optical dating (optically stimulated luminescence (OSL) and thermally-transferred OSL (TT-OSL)) of quartz grains and radiocarbon (<sup>14</sup>C) dating of hearth charcoal and avian eggshells were used to establish the occupation chronology (Fig. 2). The oldest layer (SU4) yielded calibrated <sup>14</sup>C ages of deposition of >46.0 and >44.7 ka (mean calibrated ages ± 68% probability ranges = 48.2 ± 1.2 and 47.3 ± 1.5 kyr) based on an eggshell of a large extinct megapode bird<sup>6</sup> (see Supplementary Information: megafauna), and a <sup>14</sup>C-calibrated age of 49.2–46.3 ka, based on an emu eggshell (Fig. 2).

All are within 2 standard errors of two associated optical ages of 43.8 ± 3.4 ka and 42.8 ± 2.4 ka (Extended Data Fig. 6). The age of the overlying SU3 is constrained by five <sup>14</sup>C ages with a combined span of 41.0–32.7 ka and optical ages of 40.5 ± 2.2 and 30.3 ± 1.6 ka. Deposition of SU2 is constrained to 29.8–24.4 ka from Bayesian modelling of one optical age and four <sup>14</sup>C ages, while three optical ages reveal that SU1B accumulated 11.8–9.9 ka, with post-depositional bioturbation resulting in younger eggshells being incorporated into this unit (see Extended Data Fig. 7 and Supplementary Information: chronology). These ages isolate Warraty rock shelter as the only site outside of tropical northern Australia with a rich, stratified record of repeated human activity spanning 50–10 ka.

Stone artefacts were found throughout the deposit (Fig. 2) with concentrations at depths of 5–20 cm (corresponding to SU1–upper part SU2) and 60–80 cm (SU3). The lithic assemblage is mostly composed of whole flakes, broken flakes and waste material. Stone artefacts were made from a range of raw materials, reflecting a change in use of preferred rock-type over time. Tools in the lowermost units were predominantly made of quartz. In the upper part of SU2 and in SU1, chert and silcrete become major components. This pattern was also reflected by changes in tool types; SU1 and SU2 contained predominately backed and small hafted tools, whereas SU3 and SU4 contained whole and retouched flakes (Extended Data Fig. 5). The chronology revealed by Bayesian modelling of all stratigraphically reliable ages available for Warraty greatly extends the antiquity of backed and hafted tools in Australia. The oldest backed artefacts, three geometric microliths found in SU2 at a depth of 25 cm (Fig. 2), have a Bayesian modelled age of deposition of 30–24 ka. The previous oldest confirmed deposition ages are 4 ka in the arid zone<sup>10</sup> and 8.5 ka on the east coast<sup>7</sup>. A possible occurrence of 15 ka has been reported from rock-shelter site GRE8 in the Carpentarian Gorges, northern Queensland (Fig. 1), but interpretation of its stratigraphic context has been questioned<sup>10,11</sup>. Residue analysis identified resin within SU3 (70–75 cm depth), which shows that some flakes had been hafted (see Extended Data Fig. 8A and Supplementary Information: residues). Our modelling produced a deposition age of 40–33 ka for this unit, showing that this is by far the earliest-known evidence of hafting technology in Australia and Southeast Asia. The previous oldest ages were early Holocene (10–9 ka)<sup>12,13</sup>.

White spheroids (diameter 2–30 mm) were found at depths of 20–75 cm in SU2–SU3 (40–24 ka). X-ray diffraction analysis identified the material in the spheres as gypsum (see Extended Data Fig. 8B and Supplementary Information: gypsum), the closest known source of which is 12–15 km to the north. We interpret this gypsiferous material as having been brought to the rock shelter for making white pigment. X-ray diffraction was also used to confirm the presence of red ochre,

<sup>1</sup>Department of Archaeology and History, La Trobe University, Melbourne, Victoria 3083, Australia. <sup>2</sup>Geomorphologic Consultant Gladesville, Sydney 2111, New South Wales, Australia. <sup>3</sup>School of Physical Sciences, the Environment Institute and the Institute for Photonics and Advanced Sensing, University of Adelaide, Adelaide, South Australia 5005, Australia. <sup>4</sup>School of Biological Sciences, Flinders University, Adelaide, South Australia 5001, Australia. <sup>5</sup>Institute for Photonics and Advanced Sensing, School of Physical Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia. <sup>6</sup>Defence Science and Technology Group, Edinburgh, Adelaide, South Australia 5111, Australia. <sup>7</sup>Australian Nuclear Science and Technology Organisation, Lucas Heights, Sydney, New South Wales 2234, Australia. <sup>8</sup>In the Groove Analysis Pty Ltd, Indooroopilly, Brisbane, Queensland 4068, Australia. <sup>9</sup>School of Social Sciences, University of Queensland, St Lucia, Brisbane, Queensland 4072, Australia. <sup>10</sup>Adnyamathanha Traditional Lands Association, Port Augusta, South Australia 5700, Australia.



**Figure 1 | Site distribution map showing archaeological and megafaunal sites.** Site distribution map with sea level at the  $-60$  m contour showing the geographical location of known Australian arid zone archaeological and megafaunal sites with age ranges discussed in the text. Arid interior sites 1–9, arid west coast sites 10 and 11, northern desert margins sites 13–16, southeast desert margins sites 17–21. 1, Warraty rock shelter, 49–45 kyr in age; 2, Djadjiling and Newman P2055, 41–30 kyr in age; 3, Allens Cave, 40 kyr in age; 4, Lake Gregory, >37 kyr in age; 5, Juukan 1, 37 kyr in age; 6, Puritjarra, 35 kyr in age; 7, Kulpi Mara, 34–29 kyr in age; 8, Lake Yantara, 31 kyr in age; 9, Serpents Glen, 28 kyr in age; 10, Northwest Cape sites (Jansz, Mandu Mandu, C99 and Pilgonaman),

40–24 kyr in age; 11, Silver Dollar, 35–22 kyr in age; 12, Noala, 31 kyr in age; 13, Carpenters Gap 1, 48–44 kyr in age; 14, Riwi, 45–34 kyr in age; 15, GRE8, 39–31 kyr in age; 16, Colless Creek, 30 kyr in age; 17, Lake Mungo and Willandra Lakes sites, 50–40 kyr in age; 18, Lake Menindee, 45 kyr in age; 19, Dempseys Lake, 44 kyr in age; 20, Cuddie Springs, 39–35 kyr in age; 21, Lake Tandou, 41–31 kyr in age. Archaeological sites are either open sites or rock shelters mapped within the modern Australian arid zone boundary and the Tropic of Capricorn. Megafaunal sites, black squares; rock shelters, circles; open sites, asterisks. The map also shows the general location of Warraty rock shelter (triangle symbol) (from ref. 10).

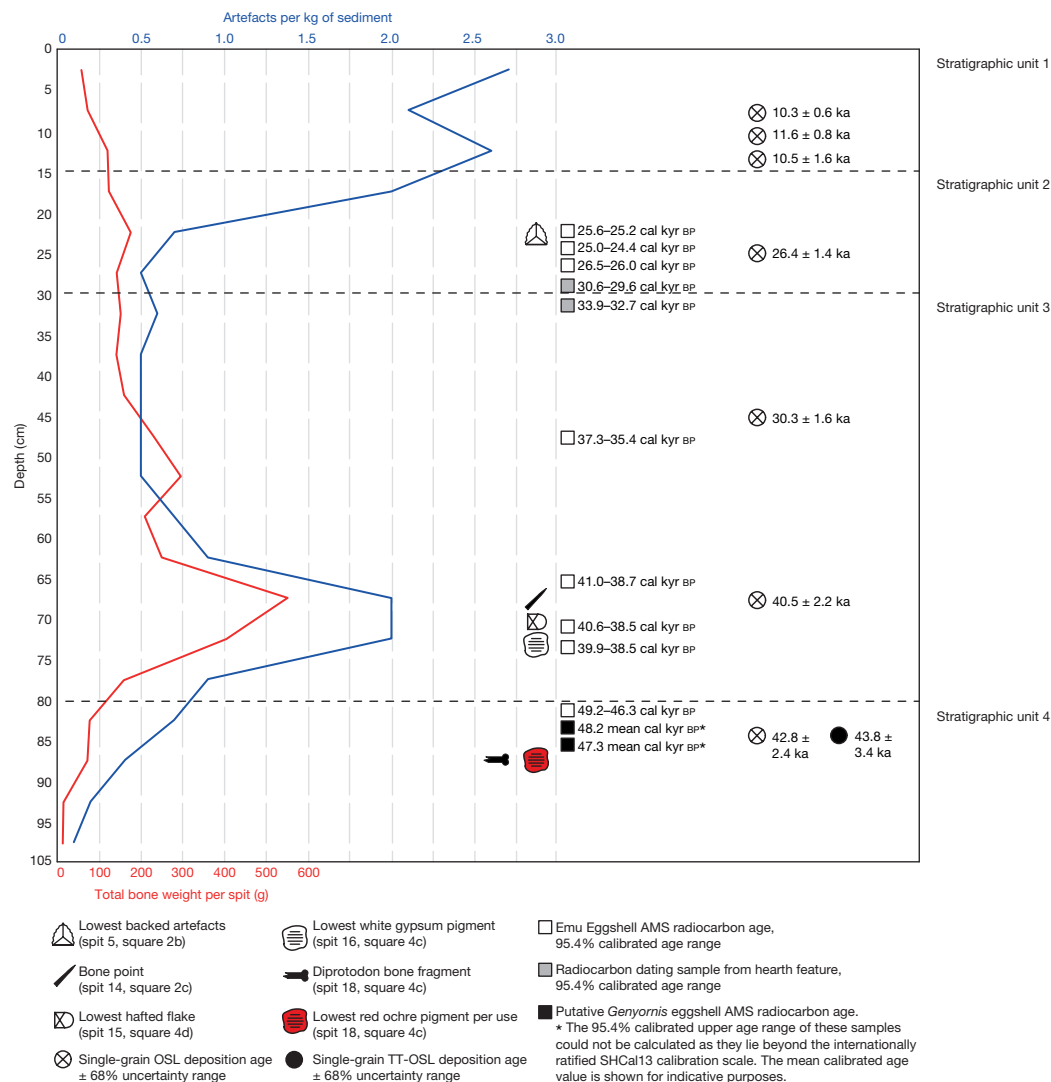
with the lowest sample found in SU4 in association with artefacts and bone material ( $\geq 49$ –46 kyr in age; Fig. 2 and Supplementary Information: red ochre). Residue analysis confirmed the presence of worked red ochre on a silcrete stone tool in SU4 at a depth of 90 cm (see Supplementary Information: residues). The earliest previous evidence for the use of ochre in Australia and Southeast Asia is 42.8 ka at Carpenters Gap rock shelter<sup>14,15</sup>. No archaeological evidence for the pre-modern use of gypsum has hitherto been reported from Australia.

At least 16 species of mammal and 1 reptile were identified from a representative sub-sample of the Warraty bone assemblage (see Supplementary Information: fauna). More than 2,000 fragments were assessed, revealing a predominance of medium-sized macropodids. Bones lacked evidence of animal gnawing or breakage patterns caused by predators or scavengers, supporting the interpretation that they accumulated as a result of human activity. A sharpened bone point, ground from the cylindrical portion of the proximal end of a macropodid fibula (similar to a yellow-footed rock-wallaby, *Petrogale xanthopus*), was recovered from a depth of 65–70 cm (SU3, Figs 2, 3). Similar single-point tools have been interpreted elsewhere as having

been used for fine needle or awl work on animal skins<sup>16</sup>. Bone tools have been considered a late Pleistocene innovation for humans in Australia and East Timor, but only appeared in the last 11 kyr in the rest of Southeast Asia<sup>17,18</sup>. The stratigraphic position of this tool indicates an age of >38 kyr, which is substantially older than the next youngest examples found at Wareen Cave (29 kyr in age) in Tasmania<sup>19</sup> and Devils Lair (26 kyr in age) in south-western Australia<sup>20,21</sup>.

A partial right juvenile radius of a rhino-sized marsupial herbivore *D. optatum* and possibly burnt and unburnt fragments of an egg-shell of a large, ground-nesting megapode bird<sup>6</sup> (shell type formerly identified as *Genyornis newtoni*; see refs 22–25) were recovered from a depth of 85–90 cm (Extended Data Figs 9, 10). Direct <sup>14</sup>C dating of the shell and optical dating of host sediments indicated a deposition age of  $\geq 49$ –46 ka (SU4, Fig. 2). The age of these fossils, together with the absence of carnivore tooth marks and the position of the shelter on a steep escarpment unsuitable for climbing by individuals of *D. optatum* (see Supplementary Information: stratigraphy), indicate co-occurrence of these taxa with humans who were probably involved in the accumulation of their remains.





**Figure 2 | Main chrono-stratigraphic and archaeological features found in Warraty rock shelter.** A graphical depiction of the distribution of artefacts and bone against depth below the surface, and the inferred stratigraphic units (1–4) in Warraty rock shelter showing overall site chronology and important artifacts found at the site. The depth of dated

samples and their ages are indicated. The density of artefacts (number per kilogram of sediment) is shown in blue and weight (g) of bone in red. White squares, emu eggshell ages; black squares, putative *Genyornis* oological material eggshell ages; white and black circles, single-grain OSL and TT-OSL ages, respectively. AMS, accelerator mass spectrometry.

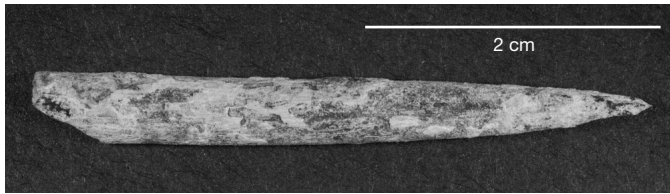
Warraty rock shelter is the oldest Australian arid-zone occupation site and one of the earliest on the continent. The presence of people in the southern interior of the continent  $\geq 49$ –46 ka (Fig. 1) suggests that, following their arrival in Australia, people dispersed more rapidly across the continent than previously thought. The location of Warraty could imply a more direct north–south route for pioneering human settlers rather than an exclusive coastal route. The evidence supports the model that Aboriginal people had settled in the Australian arid zone well before the extreme arid conditions of the last glacial maximum and the associated expansion of major environmental barriers such as sandy deserts<sup>9</sup>.

Human occupation was repeated but ephemeral in nature, indicating that Aboriginal people may have used Warraty both as a refuge at a time when the surrounding lowlands and open plains were too arid to exploit and as a temporary campsite when environmental conditions became more stable regionally<sup>10</sup>.

The development of worked-bone technology by at least 40–38 ka, hafted tools by at least 35 ka and backed artefacts by at least 24 ka shows that people living at Warraty were early innovators of modern technological adaptations found in late Pleistocene Australia and Southeast Asia. This refutes previously held views concerning the timing of cultural and technological innovation for late Pleistocene Australia<sup>8</sup>.

Warraty also provides stratified archaeological data and a chronology that directly link humans with megafauna in Australia. The late Pleistocene is marked by the extinction of large vertebrates from the continents<sup>22–24</sup>. Although at least 22 species, which later went extinct, overlapped temporally with humans in Australia and New Guinea<sup>22</sup>, only two sites, Cuddie Springs in eastern Australia and Nombe rock shelter in New Guinea (Fig. 1), have been reported to contain cultural and megafaunal material within the same stratigraphic layers<sup>22,24</sup>. However, the evidence of direct association between megafauna and humans at these sites has been challenged based on site formation, climatic, stratigraphic and chronological grounds<sup>26–28</sup>.

The discovery of megafaunal bone and directly dated eggshell in a well-stratified and reliably dated archaeological context at Warraty not only shows that these taxa were contemporary with humans but also provides the only direct evidence that people interacted with some megafauna in Australia. The location of Warraty in northern South Australia is also important for evaluating the causes of continent-wide megafaunal extinction, because it confirms the temporal overlap of humans and extinct species 50–40 ka across a much broader geographic area of Australia than previously thought. Until now, direct evidence for the co-existence of humans and megafauna had been lacking for the arid interior, a major region of the continent.



**Figure 3 | Earliest bone point.** A sharpened bone point, likely to be ground from the cylindrical portion of the proximal end of a macropod fibula similar in size to that of *Petrogale xanthopus*, the yellow-footed rock-wallaby. This specimen was found in SU3, dated to 40–38 ka and is currently the oldest bone point in Australia.

Archaeological sites with evidence of modern human colonisation, unique cultural innovation and interaction with now-extinct megafauna are rare in southern Asia and Australia. Sites preserving 50-kyr records of human occupation are rarer still. In addition to these landmark discoveries, Warraty rock shelter reveals evidence for the development of modern human behaviour in Australia and Asia. Important technological innovations and early symbolic behaviour<sup>9,29</sup> reveal that a dynamic, adaptive Aboriginal culture existed in arid Australia within only a few millennia of settlement on the continent.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 26 May; accepted 27 September 2016.**

**Published online 2 November 2016.**

- Veth, P. & O'Connor, S. in *Cambridge History of Australia, Volume 1 Colonial Australia* (eds Bashford, A. & Macintyre, S.) 1–34 (Cambridge Univ. Press, 2013).
- Allen, J. & O'Connell, J. Both half right: updating the evidence for dating first human arrivals in Sahul. *Aust. Archaeol.* **79**, 86–108 (2014).
- Mellars, P. Going east: new genetic and archaeological perspectives on the modern human colonization of Eurasia. *Science* **313**, 796–800 (2006).
- Boivin, N., Fuller, D. Q., Dennel, R., Allaby, R. & Petraglia, M. D. Human dispersal across diverse environments of Asia during the Upper Pleistocene. *Quat. Int.* **300**, 32–47 (2013).
- Roberts, R. G. *et al.* New ages for the last Australian megafauna: continent-wide extinction about 46,000 years ago. *Science* **292**, 1888–1892 (2001).
- Grellet-Tinner, G., Spooner, N. A. & Worthy, T. H. Is the 'Genyornis' egg of mihring or another extinct bird from the Australian dreamtime? *Quat. Sci. Rev.* **133**, 147–164 (2016).
- Attenbrow, V., Robertson, G. & Hiscock, P. The changing abundance of backed artefacts in south-eastern Australia: a response to Holocene climate change? *J. Archaeol. Sci.* **36**, 2765–2770 (2009).
- Habgood, P. J. & Franklin, N. R. The revolution that didn't arrive: a review of Pleistocene Sahul. *J. Hum. Evol.* **55**, 187–222 (2008).
- Hiscock, P. & Wallis, L. in *Desert Peoples: Archaeological Perspectives* (eds Veth, P., Smith, M. & Hiscock, P.) 34–57 (Blackwell Publishers, 2005).
- Smith, M. *The Archaeology of Australia's Deserts* (Cambridge Univ. Press, 2013) 185–187.
- Slack, M. J., Fullagar, R. L. K., Field, J. H. & Border, A. New Pleistocene ages for backed artefact technology in Australia. *Archaeol. in Oceania* **39**, 131–137 (2004).
- Mulvaney, D. J. & Kamminga, J. *The Prehistory of Australia* (Allen & Unwin Pty Ltd, 1999).
- Barton, H. M., Piper, P. J., Rabett, R. & Reeds, I. Composite hunting technologies from the terminal Pleistocene and early Holocene, Niah Cave Borneo. *J. Archaeol. Sci.* **36**, 1708–1714 (2009).
- O'Connor, S. & Frankhauser, B. in *Histories of Old Ages: Essays in Honour of Rhys Jones* (eds Anderson, A., Lilley, I. & O'Connor, S.) 287–300 (Pandan Books, Research School of Pacific and Asian Studies, Australian National Univ., 2001).
- Aubert, M. *et al.* Pleistocene cave art from Sulawesi, Indonesia. *Nature* **514**, 223–227 (2014).

- Webb, C. & Allen, J. A functional analysis of Pleistocene bone tools from two sites in southwest Tasmania. *Archaeol. in Oceania* **25**, 75–78 (1990).
- O'Connor, S., Robertson, G. & Aplin, K. P. Are osseous artefacts a window to perishable material culture? Implications of an unusually complex bone tool from the Late Pleistocene of East Timor. *J. Hum. Evol.* **67**, 108–119 (2014).
- Rabett, J. R. The early exploitation of South-east Asian mangroves: bone technology from caves and open sites. *Asian Perspect.* **44**, 154–179 (2005).
- Cosgrove, R. Forty two degrees south: the archaeology of Late Pleistocene Tasmania. *J. World Prehist.* **13**, 357–402 (1999).
- Dortch, C. E. *Devil's Lair, A Study in Prehistory* (Western Australian Museum, 1984).
- Dortch, C. E. & Dortch, J. Review of Devil's Lair artefact classification and radiocarbon chronology. *Aust. Archaeol.* **43**, 28–32 (1996).
- Barnosky, A. D., Koch, P. L., Feranec, R. S., Wing, S. L. & Shabel, A. B. Assessing the causes of late Pleistocene extinctions on the continents. *Science* **306**, 70–75 (2004).
- Field, J. H., Wroe, S., Trueman, C. N., Garvey, J. & Wyatt-Spratt, S. Looking for the archaeological signature in Australian megafaunal extinctions. *Quat. Int.* **285**, 76–88 (2013).
- Wroe, S. *et al.* Climate change frames debate over the extinction of megafauna in Sahul (Pleistocene Australia–New Guinea). *Proc. Natl Acad. Sci. USA* **110**, 8777–8781 (2013).
- Miller, G. H. *et al.* Pleistocene extinction of *Genyornis newtoni*: human impact on Australian megafauna. *Science* **283**, 205–208 (1999).
- Brook, B. W. *et al.* Lack of chronological support for stepwise prehuman extinctions of Australian megafauna. *Proc. Natl Acad. Sci. USA* **110**, E3368 (2013).
- Cohen, T. *et al.* Hydrological transformation coincided with megafaunal extinction in central Australia. *Geology* **43**, 195–198 (2015).
- Grün, R. *et al.* ESR and U-series analyses of faunal material from Cuddie Springs, NSW, Australia: implications for the timing of the extinction of the Australian megafauna. *Quat. Sci. Rev.* **29**, 596–610 (2010).
- Balme, J., Davidson, I., McDonald, J., Stern, N. & Veth, P. Symbolic behaviour and the peopling of the southern arc route to Australia. *Quat. Int.* **202**, 59–68 (2009).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** The fieldwork was undertaken with the approval of Adnyamathanha Traditional Lands Association (ATLA) and the South Australian Department of Aboriginal Affairs and Reconciliation. We thank for their contributions: A. Coulthard who participated in the fieldwork; R. Frank for his support both in the field and with drawing site maps; field assistants L. Foley and S. Adams; R. Cosgrove provided faunal advice; G. Robertson assisted with residue laboratory work; C. Brown assisted with graphic design, N. Bonney provided botanical advice; M. Raven and J. Webb carried out analyses of ochres; G. Medlin provided faunal advice; B. Barker provided botanical advice; P. Toms provided access to single grain OSL dating assessment; F. Williams for assistance in fieldwork and sample collection for optical dating; A. Couzens, W. Handley and G. Gully for their assistance with micro-CT imaging and photography of megafaunal specimens; and P. Veth, P. Hiscock, B. Cundy, W. Shawcross, S. Webb, J. Magee and D. Witter for critical feedback on the manuscript. This research was supported by Alinta Energy who provided financial support for fieldwork carried out in May 2013. L.A. and G.P. were supported by Australian Research Council Future Fellowship Grants FT130100195 and FT130101728, respectively. V.L. and Centre for Accelerator Science at ANSTO, acknowledge the support of the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS). FWS is who - they did the drawing of the lithics and ought to be here.

**Author Contributions** G.H. conceived the project, supervised fieldwork, undertook data collection and analysis. L.A. provided chronological assessment for sediments and undertook the Bayesian age modelling. G.P. and T.W. identified megafaunal remains. D.Q. prepared chronological samples. N.S. advised on the results of chronological assessments. V.L. assisted with radiocarbon dating issues. E.F. provided fieldwork support and assisted in the analysis of stone artefact material. B.S. analysed and reported on stone artefact residues. V.C. and C.C. provided cultural advice and logistical support. C.C., S.W. and D.J. assisted with fieldwork. G.H., P.M., L.A., G.P. and T.W. wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.H. (A.R.A.S@bigpond.com).

## METHODS

**Excavation methods.** Hand excavation was undertaken within Warraty rock shelter (Extended Data Fig. 1a–c), using two  $2 \times 1$  m trenches (Extended Data Fig. 1b, c). The first exploratory trench (squares 4C and 4D) was located on the southern side of the shelter to provide a cross-section through potential living areas from east to west (taking in part of the back wall). The second trench (squares 2B and 2C) was positioned in the centre of the shelter floor. Square 2B was only excavated down to spit 6 (30 cm) until obstructed by a large roof fall block.

Excavation was carried out using a 1-m grid system. This grid was further subdivided into 25 cm square units (quadrats: A–D Extended Data Fig. 1c) to enable greater recording precision. Excavation was undertaken using 5 cm spits, with a trowel and hand shovel and the excavated material was dry-sieved through (8, 5 and 2 mm) sieves. Stratigraphic features, such as charcoal lenses, ash or hearth-like lenses, were all drawn, photographed and noted during the excavation process. Detailed profile drawings were made on completion of the excavation. If possible, stone artefacts, bone and ochre were individually recorded and excavated from their *in situ* location and bagged separately, rather than being retrieved from the bulk sediment collected from individual spits. To reduce the potential for contamination of the excavated surfaces by loose material falling into the pit, the area surrounding each square (where possible) was covered with plastic sheeting and any loose spoil removed. Excavators were encouraged to wear soft shoes to reduce potential damage to the pit walls.

Each spit was recorded using a standard excavation sheet, which included making notes of the individual features and conditions encountered within each square and quadrat. A geomorphologist (P.M.) was present during the excavation process to advise excavators on the variability of sedimentary units and the positions of potential depositional changes associated with strata boundaries. When these depositional changes were detected during the excavation, they were noted and new strata were excavated as separate entities within a spit. This process reduced the potential for the sediment of two different strata being mixed.

Where a filled burrow (mostly attributable to rabbits) was encountered, the infilling sediment was removed as a discrete unit and bagged separately. These burrows were excavated in each square before the undisturbed enclosing sediments were excavated, thereby minimising potential contamination of the primary sediment by younger and/or mixed burrow infill material.

All *in situ* finds larger than 2 cm (including bone) were plotted in plan view within each excavated quadrat (A–D) and their depth ascertained by levelling within each spit (Extended Data Fig. 1b, c). Some of these *in situ* finds were photographed, however, most were not as their importance at the time of the excavation was not known (for example, bone point and *Diprotodon* bone fragments). All sediment removed from each spit was weighed. All material remaining on the sieves was bagged for further analysis. Bulk soil samples were removed for sedimentary analyses. At a later stage, sediment residues were wet-sieved through a 1-mm mesh before sorting. All charcoal, artefacts, bone fragments and plant matter recovered from this process were bagged. The location of charcoal samples collected during the excavation was plotted on plans, their depth recorded and then bagged separately.

A total of 1,070 stone artefacts were tested for refitting from square 2C in order to confirm the integrity of the apparently intact depositional laminae. These artefacts were assessed to see if they could be refitted together as parts of a former piece of stone artefact, on the assumption that their original separation was attributable to stone reduction activity and that they must have originated on a single surface in the shelter. All artefacts from square 2C that were greater than 10 mm in maximum dimension were assessed for refitting. Each refit set comprised artefacts that could be treated analytically as a single knapping episode<sup>30</sup>. As such, the distance between artefact elevations for each refitting set provides a proxy measure for the vertical displacement of cultural material through the deposit by human trampling and other post-depositional activity<sup>31</sup>. Refitting was attempted within a  $1 \times 1$  m area, which represents only a sample of the occupied area and therefore is unlikely to capture an entire knapping event. In addition, artefacts will have undoubtedly been laterally dispersed as a result of repeated human and animal occupation; therefore we can predict that the overall percentage of refits (the ‘success rate’) will be low<sup>32</sup>.

**Optical dating of quartz grains.** Seven optical dating samples were collected from cleaned exposures of excavation square 2C using metal tubes, and wrapped in light-proof bags for transportation and storage. Bulk sediment samples were also collected from the surrounding few centimetres of each sample tube for  $\beta$ -dose rate determination and water content analysis. In the laboratory, quartz grains of 180–212  $\mu$ m diameter were extracted from the un-illuminated centres of the metal tubes under safe light (dim red LED) conditions and prepared for burial dose estimation using standard procedures<sup>33</sup>, including etching by 48% hydrofluoric acid for 40 min to remove the  $\alpha$ -irradiated external layers.

Two semi-independent approaches were used to obtain optical age estimates for the Warraty samples. Single-grain optically stimulated luminescence (OSL) dating

of quartz<sup>34,35</sup> was routinely applied to all samples, and was preferred over standard multiple-grain OSL dating because of its ability to identify insufficiently bleached grain populations<sup>36,37</sup>, contaminant grains associated with post-depositional mixing<sup>38</sup>, and aberrant grains displaying inherently unsuitable luminescence properties<sup>39,40</sup>. Single-grain TT-OSL dating of quartz<sup>41–43</sup> was applied to the oldest sample in the sequence (ERS-7) as a means of cross-checking the reliability of the OSL-dating approach over dose ranges of 150–200 Gy. TT-OSL dating was applied to individual grains of quartz rather than multi-grain aliquots in this study, following the reliable application of this approach at other archaeological cave and rock shelter sites<sup>38,44,45</sup>.

OSL and TT-OSL equivalent dose ( $D_e$ ) measurements were made using experimental apparatus and quality assurance criteria described previously<sup>38,44</sup>. Samples were irradiated with a Risø TL-DA-20 <sup>90</sup>Sr/<sup>90</sup>Y  $\beta$  source that had been calibrated to administer known doses to multi-grain aliquots and single-grain discs. For single-grain measurements, spatial variations in the  $\beta$ -dose rate across the disc plane were taken into account by undertaking hole-specific calibrations using  $\gamma$ -irradiated quartz. Quartz grains with a diameter of 180–212  $\mu$ m were measured in standard single-grain aluminium discs drilled with an array of  $300 \times 300$   $\mu$ m holes.

$D_e$  values were determined for individual grains of quartz using the single-aliquot regenerative-dose (SAR) procedures<sup>34</sup> shown in Supplementary Table 3. Between 900 and 1,300 individual quartz grains of each sample were measured for  $D_e$  determination (Supplementary Table 4). Sensitivity-corrected dose–response curves were constructed using the first 0.17 s of each green-laser stimulation after subtracting a mean background count obtained from the last 0.25 s of the signal. A preheat of 260 °C for 10 s was used in the OSL SAR procedure before measuring the natural ( $L_n$ ) and regenerative dose ( $L_x$ ) signals, and a cut-heat of 160 °C was applied before undertaking the test-dose OSL measurements ( $T_n$  and  $T_x$ ) (Supplementary Table 3). These preheating conditions yielded an accurate measured-to-recovered dose ratio of  $1.03 \pm 0.03$  and a relatively low overdispersion value of  $12 \pm 3\%$  for a  $\sim 130$  Gy dose-recovery test performed on individual grains of sample ERS-5 (Extended Data Fig. 6A). The single-grain TT-OSL SAR procedure (Supplementary Table 3) uses a TT-OSL test dose measurement rather than an OSL test-dose measurement (step 11) to correct for sensitivity change, following suitability assessments performed elsewhere<sup>40,44–46</sup>. A dose-recovery test performed on 1,200 individual quartz grains of sample ERS-7 attests to the general suitability of this SAR procedure (Extended Data Fig. 6A). The TT-OSL dose-recovery test was performed on a batch of unbleached grains owing to the relatively long periods of light exposure needed to bleach natural TT-OSL signals down to low residual levels<sup>47</sup>. A known (172 Gy) laboratory dose of similar magnitude to the expected  $D_e$  was added on top of the natural signals for these grains. The recovered dose was then calculated by subtracting the weighted mean natural  $D_e$  of sample ERS-7 ( $168 \pm 12$  Gy; determined on a separate batch of grains and summarized in Supplementary Table 5) from the weighted mean  $D_e$  of these unbleached and dosed grains ( $350 \pm 18$  Gy). This approach yielded a net (that is, natural-subtracted) recovered-to-given ratio of  $1.06 \pm 0.09$  for sample ERS-7. An overdispersion value of  $12 \pm 9\%$  was calculated for the  $D_e$  distribution of the unbleached and dosed batch of grains, which is consistent with that obtained for the single-grain OSL dose-recovery test.

Individual  $D_e$  estimates are presented with their 1 standard error ranges (Supplementary Table 5 and Extended Data Fig. 6C), which have been derived from three sources of uncertainty: (i) a random uncertainty term arising from photon-counting statistics for each OSL or TT-OSL measurement, calculated using equation 3 from ref. 48; (ii) an empirically determined instrument-reproducibility uncertainty of 2% for each single-grain measurement; and (iii) a dose–response curve fitting uncertainty determined using 1,000 iterations of the Monte Carlo method implemented in Analyst<sup>49</sup>.

Environmental dose rates have been calculated using a combination of *in situ* field  $\gamma$ -ray spectrometry and high-resolution  $\gamma$  spectrometry of dried and homogenized bulk sediments collected directly from the OSL-sampling positions. Cosmic-ray dose-rate contributions were calculated using the equations in ref. 50 after taking into consideration site altitude, geomagnetic latitude, density, thickness and geometry of the sediment or bedrock overburden. A small, assumed internal ( $\alpha$  plus  $\beta$ ) dose rate of  $0.03 \pm 0.01$  Gy per kyr has been included in the final dose-rate calculations, based on published <sup>238</sup>U and <sup>232</sup>Th measurements for etched quartz grains from a range of locations<sup>51–54</sup> and an  $\alpha$  efficiency factor ( $a$  value) of  $0.04 \pm 0.01$  (refs 55–57). Radionuclide concentrations and specific activities have been converted to dose rates using published conversion factors<sup>58</sup>, allowing for  $\beta$ -dose attenuation<sup>59,60</sup> where applicable.

**Radiocarbon (<sup>14</sup>C) dating of hearth charcoal and eggshell samples.** Seventeen eggshell fragments (15 Dromaius, 2 megapode) and two large charcoal fragments associated with hearth features were used to derive the final <sup>14</sup>C chronology of SU1A to SU4. A range of other organic materials were initially submitted for <sup>14</sup>C analysis but these were considered unreliable based on stratigraphic or



methodological grounds, as detailed in Supplementary Information (radiocarbon dating). Samples were pretreated using acid–base–acid procedures and their  $^{14}\text{C}$  contents were measured using accelerator mass spectrometry at the Waikato Radiocarbon Laboratory and the ANSTO Radiocarbon Facility. Uncalibrated eggshell and charcoal  $^{14}\text{C}$  ages (Supplementary Table 7) are expressed in  $^{14}\text{C}$  years before present ( $^{14}\text{C}$  yr BP, where BP is defined as 1950 AD) following standard reporting conventions<sup>61</sup>. Isotopic fractionation has been corrected for by using the measured  $\delta^{13}\text{C}$  value of each sample. The  $^{14}\text{C}$  age estimates have been calibrated with the internationally ratified southern hemisphere SHCal13 curve<sup>62</sup>, using OxCal v4.2.4 (ref. 63). The calibrated  $^{14}\text{C}$  age ranges (cal yr BP) are described as 95.4% probability ranges throughout.

**Bayesian age modelling.** Bayesian age modelling was used to integrate all stratigraphically reliable chronological information within a unified statistical framework and to derive combined age estimates for individual stratigraphic units. The Bayesian age model for Warratyti was constructed using OxCal v4.2.4 (ref. 63). A sequence-deposition model with nested phases of uniform prior duration and associated boundaries was used to derive a combined chronostratigraphic framework for the site, following the approach outlined previously<sup>64</sup>. The main depositional sequence incorporates the five stratigraphic units found at Warratyti (SU1A–SU4) in an ordered succession according to depth. Units have been represented as nested sequences within the broader depositional column as their overall stratigraphic ordering is sufficiently well preserved. Separate phases have been used to represent groupings of numerical ages within individual units. This approach was deemed necessary for SU2–SU4 because the  $^{14}\text{C}$  and optical dating samples were collected from multiple excavation squares that exhibited potentially irregular or spatially heterogeneous stratigraphic relationships. The exact relative ordering of dating samples from each unit could therefore not be directly constrained in a vertical profile. SU1A and SU1B are internally heterogeneous and exhibit signs of inter-horizon mixing, as borne out by the multiple dose components of optical dating samples ERS-1, ERS-2 and ERS-3 (see Supplementary Information: chronology). This has prevented us from making any assumptions about relative chronological ordering of dated horizons within these units, although the relative ordering of SU1A and SU1B remains sufficiently clear. Groupings within SU1A and SU1B were therefore also nested as separate phases rather than sub-sequences.

Each stratigraphic unit has been represented by a single phase, with the exception of SU3. This unit is significantly thicker than the others at Warratyti and can be differentiated into two broad archaeo–palaeontological phases according to distinct changes in artefact content and bone abundance at a depth of approximately 60–65 cm (Fig. 2). Two separate phases were nested within the SU3 sequence to account for archaeological sub-structuring of the dated horizons. Boundaries were used to delineate the beginning and end of each stratigraphic unit. We have not, however, incorporated any prior depositional gaps in the sequence model as there is no direct evidence of unconformities at Warratyti. The entire site sequence has been constrained with a minimum age of 0 years before 1950 AD and a maximum age of 60 kyr before 1950 AD. The latter represents a conservative upper age estimate for this archaeological sequence and has been chosen to predate sufficiently the earliest existing evidence of human presence in Australia around  $\leq 50$  ka, as determined from an assessment of 26 early occupation sites across Sahul<sup>65</sup>. We note, however, that the Bayesian age model is largely insensitive to our choice of maximum age constraint given the range of likelihoods obtained on the lowest stratigraphic unit.

The  $^{14}\text{C}$  data were input into the model as conventional ages (using the OxCal R\_Date function) and were subsequently calibrated using the SHCal13 curve<sup>62</sup> as part of the modelling procedure. Modelled posterior  $^{14}\text{C}$  age ranges are therefore presented in calendar years before 1950 AD (cal yr BP). To avoid introducing systematic errors in the posterior results, optical dating ages (calculated as kyr before sample collection in 2012 AD) were similarly converted to years before 1950 AD before their incorporation in the model.

The final Bayesian age model for Warratyti was run using the general outlier function<sup>66</sup>, which is based on a Student's  $t$ -test distribution with 5 degrees of freedom. Prior outlier probabilities of 5% were equally assigned to all dating samples to identify potentially significant statistical outliers. Likelihood estimates that yielded posterior outlier probabilities  $>5\%$  were not excluded from the final model but were proportionally down-weighted in the iterative Monte Carlo runs, thereby producing an averaged chronological model<sup>66</sup>.

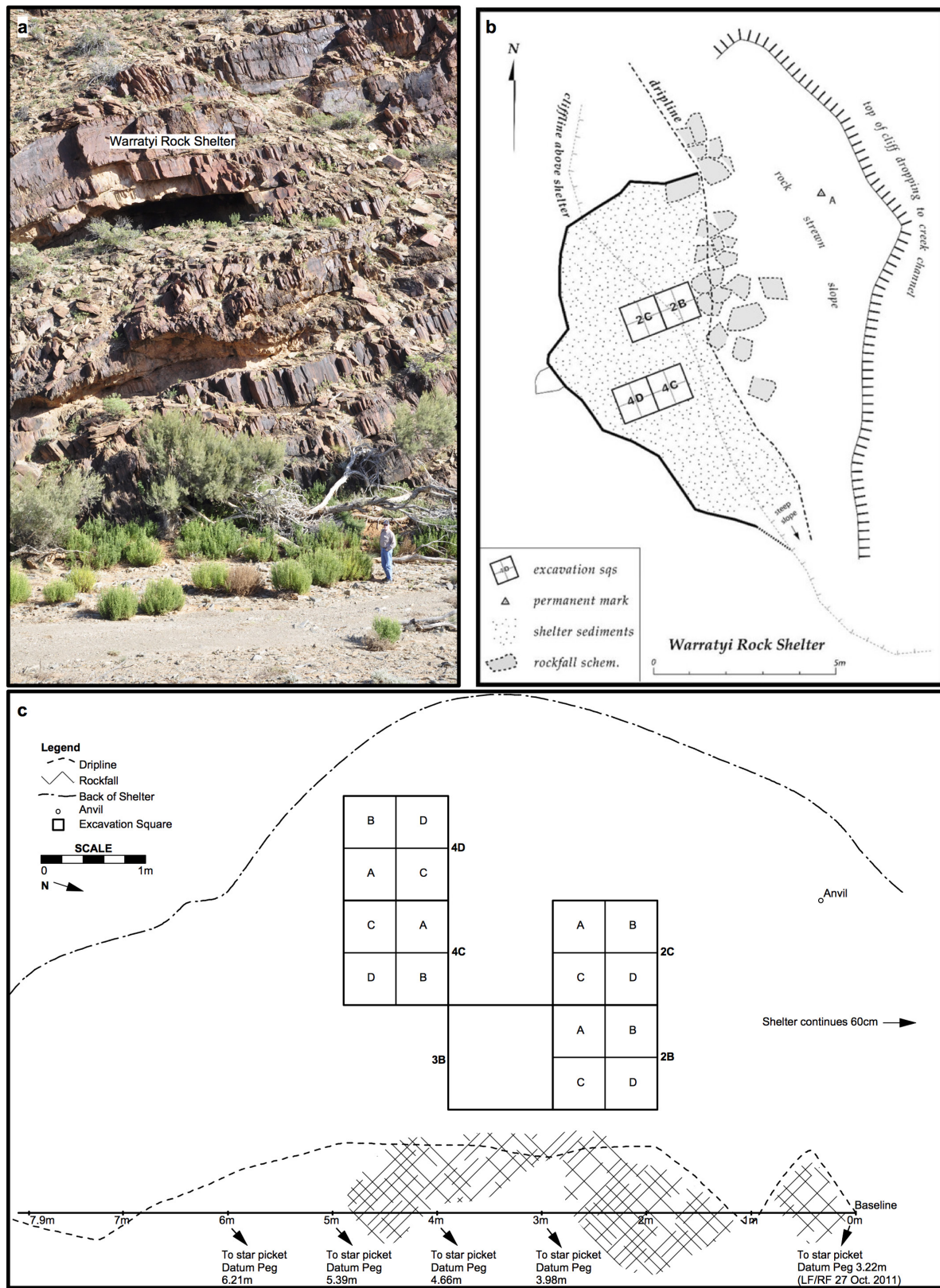
**Use-wear and residue analyses.** Use-wear studies were conducted using a hand-held polarizing Dino-Lite AM4815ZT microscope at magnifications of  $30\times$  to  $230\times$ . Additional high-power microscopic investigation using an Olympus BX51 at magnifications of  $50\times$  and  $500\times$  in brightfield and darkfield, were undertaken on seven artefacts. Within the overall study, six small samples of residue (approximately 10–20  $\mu\text{l}$ ) were extracted from margins or areas of interest on specific stone artefacts. Samples were extracted with pipettes using distilled water as the lifting

medium and transferred to slides that had been pre-cleaned with ethanol. The Dino-lite was used to guide this process. Sample slides were dried under covers for 24 h and then stained with a 0.25% solution of picosirius red using the protocols described previously<sup>67</sup>. A Leitz Dialux 22 microscope with polarizing capability was used to examine the stained slides. Residues were photographed in plane, part-polarized and cross-polarized light at a magnification of  $400\times$  using a Tucsen ISH 500 camera.

**Data availability.** All elements necessary to allow interpretation and replication of results, including full datasets and detailed experimental procedures are provided in the Supplementary Information. Fossils and archaeological material generated in this study will be deposited at the SA Museum and an Adnyamathanha Traditional Lands Association (ATLA)-keeping place within a 6–12-month time-frame and will be publicly accessible upon request with permission from ATLA and the corresponding author.

30. Larson, M. L. & Kornfeld, M. Chipped stone nodules: theory, method, and examples. *Lithic Technol.* **22**, 4–18 (1997).
31. Villa, P. Conjoinable pieces and site formation processes. *Am. Antiq.* **47**, 276–290 (1982).
32. Laughlin, J. P. & Kelly, R. L. Experimental analysis of the practical limits of lithic refitting. *J. Archaeol. Sci.* **37**, 427–433 (2010).
33. Aitken, M. J. *An Introduction to Optical Dating: the Dating of Quaternary Sediments by the Use of Photon-stimulated Luminescence* (Oxford Univ. Press 1998).
34. Murray, A. S. & Roberts, R. G. Determining the burial time of single grains of quartz using optically stimulated luminescence. *Earth Planet. Sci. Lett.* **152**, 163–180 (1997).
35. Bøtter-Jensen, L., Bulur, E., Duller, G. A. T. & Murray, A. S. Advances in luminescence instrument systems. *Radiat. Meas.* **32**, 523–528 (2000).
36. Olley, J. M., Pietsch, T. & Roberts, R. G. Optical dating of Holocene sediments from a variety of geomorphic settings using single grains of quartz. *Geomorphology* **60**, 337–358 (2004).
37. Arnold, L. J., Bailey, R. M. & Tucker, G. E. Statistical treatment of fluvial dose distributions from southern Colorado arroyo deposits. *Quat. Geochronol.* **2**, 162–167 (2007).
38. Arnold, L. J., Demuro, M. & Navazo Ruiz, M. Empirical insights into multi-grain averaging effects from 'pseudo' single-grain OSL measurements. *Radiat. Meas.* **47**, 652–658 (2012).
39. Demuro, M. *et al.* Optically stimulated luminescence dating of single and multiple grains of quartz from perennially frozen loess in western Yukon territory, Canada: comparison with radiocarbon chronologies for the late Pleistocene Dawson tephra. *Quat. Geochronol.* **3**, 346–364 (2008).
40. Demuro, M., Arnold, L. J., Froese, D. G. & Roberts, R. G. OSL dating of loess deposits bracketing Sheep Creek tephra beds, northwest Canada: dim and problematic single-grain OSL characteristics and their effect on multi-grain age estimates. *Quat. Geochronol.* **15**, 67–87 (2013).
41. Wang, X. L., Wintle, A. G. & Lu, Y. C. Thermally transferred luminescence in fine-grained quartz from Chinese loess: basic observations. *Radiat. Meas.* **41**, 649–658 (2006).
42. Arsuaga, J. L. *et al.* Neandertal roots: cranial and chronological evidence from Sima de los Huesos. *Science* **344**, 1358–1363 (2014).
43. Demuro, M., Arnold, L. J., Parés, J. M. & Sala, R. Extended-range luminescence chronologies suggest potentially complex bone accumulation histories at the Early-to-Middle Pleistocene palaeontological site of Huéscar-1 (Guadix-Baza basin, Spain). *Quat. Int.* **389**, 191–212 (2015).
44. Arnold, L. J. *et al.* Luminescence dating and palaeomagnetic age constraint on hominins from Sima de los Huesos, Atapuerca, Spain. *J. Hum. Evol.* **67**, 85–107 (2014).
45. Demuro, M. *et al.* New luminescence ages for the Galería Complex archaeological site: resolving chronological uncertainties on the acheulean record of the Sierra de Atapuerca, northern Spain. *PLoS One* **9**, e110169 (2014).
46. Stevens, T., Buylaert, J.-P. & Murray, A. S. Towards development of a broadly-applicable SAR TT-OSL dating protocol for quartz. *Radiat. Meas.* **44**, 639–645 (2009).
47. Tsukamoto, S., Duller, G. A. T. & Wintle, A. G. Characteristics of thermally transferred optically stimulated luminescence (TT-OSL) in quartz and its potential for dating sediments. *Radiat. Meas.* **43**, 1204–1218 (2008).
48. Galbraith, R. F. A note on the variance of a background-corrected OSL count. *Ancient TL* **20**, 49–51 (2002).
49. Duller, G. A. T. Assessing the error on equivalent dose estimates derived from single aliquot regenerative dose measurements. *Ancient TL* **25**, 15–24 (2007).
50. Prescott, J. R. & Hutton, J. T. Cosmic ray contributions to dose rates for luminescence and ESR dating: large depths and long-term time variations. *Radiat. Meas.* **23**, 497–500 (1994).
51. Mejdahl, V. Internal radioactivity in quartz and feldspar grains. *Ancient TL* **5**, 10–17 (1987).
52. Bowler, J. M. *et al.* New ages for human occupation and climatic change at Lake Mungo, Australia. *Nature* **421**, 837–840 (2003).
53. Jacobs, Z., Duller, G. A. T. & Wintle, A. G. Interpretation of single-grain  $D_e$  distributions and calculation of  $D_e$ . *Radiat. Meas.* **41**, 264–277 (2006).
54. Pawley, S. M. *et al.* Age limits on Middle Pleistocene glacial sediments from OSL dating, north Norfolk, UK. *Quat. Sci. Rev.* **27**, 1363–1377 (2008).

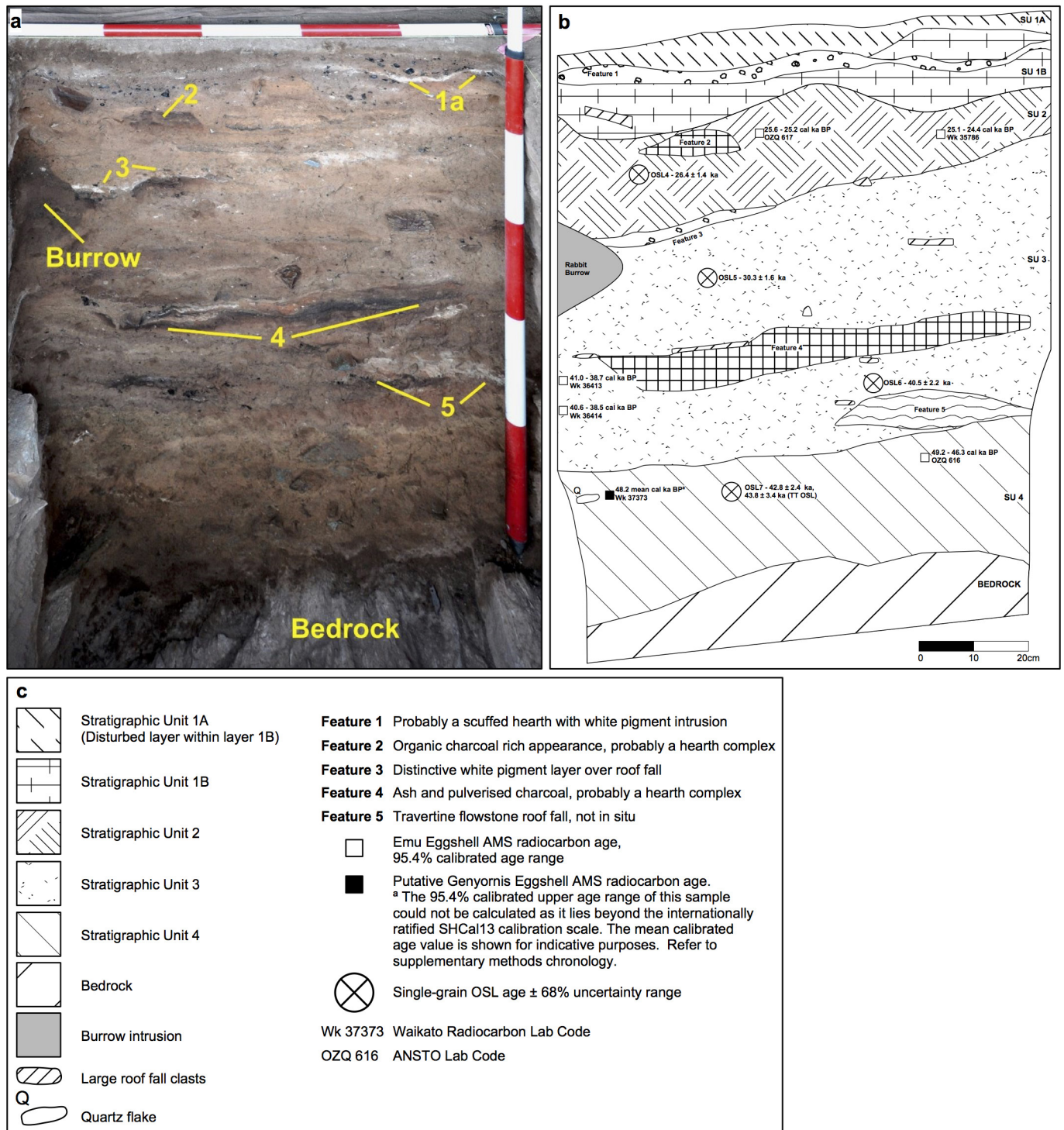
55. Questiaux, D. Optical dating of loess: comparisons between different grain size fractions for infrared and green excited wavelengths. *Nucl. Tracks Radiat. Meas.* **18**, 133–139 (1991).
56. Rees-Jones, J. Optical dating of young sediments using fine-grain quartz. *Ancient TL* **13**, 9–14 (1995).
57. Rees-Jones, J. & Tite, M. S. Optical dating results for British archaeological sediments. *Archaeometry* **39**, 177–187 (1997).
58. Guérin, G., Mercier, N. & Adamiec, G. Dose-rate conversion factors: update. *Ancient TL* **29**, 5–8 (2011).
59. Mejdahl, V. Thermoluminescence dating: beta-dose attenuation in quartz grains. *Archaeometry* **21**, 61–72 (1979).
60. Brennan, B. J. Beta doses to spherical grains. *Radiat. Meas.* **37**, 299–303 (2003).
61. Stuiver, M. & Polach, H. A. Reporting of  $^{14}\text{C}$  data. *Radiocarbon* **19**, 355–363 (1977).
62. Hogg, A. G. *et al.* SHCal13 Southern Hemisphere calibration 0–50,000 years cal bp. *Radiocarbon* **55**, 1889–1903 (2013).
63. Bronk Ramsey, C. Radiocarbon calibration and analysis of stratigraphy: the OxCal program. *Radiocarbon* **37**, 425–430 (1995).
64. Macken, A. C., Staff, R. A. & Reed, E. H. Bayesian age-depth modelling of Late Quaternary deposits from Wet and Blanche Caves, Naracoorte, South Australia: a framework for comparative faunal analyses. *Quat. Geochronol.* **17**, 26–43 (2013).
65. O'Connell, J. F. & Allen, J. The process, biotic impact, and global implications of the human colonization of Sahul about 47,000 years ago. *J. Archaeol. Sci.* **56**, 73–84 (2015).
66. Bronk Ramsey, C. Dealing with offsets and outliers in radiocarbon dating. *Radiocarbon* **51**, 1023–1045 (2009).
67. Stephenson, B. A modified Picro-Sirius Red (PSR) staining procedure with polarization microscopy for identifying collagen in archaeological residues. *J. Archaeol. Sci.* **61**, 235–243 (2015).



**Extended Data Figure 1 | Warraty rock shelter, site location, site plan and excavation strategy.** **a**, Warraty rock shelter above a dry creek bed, looking west at the main ridge profile and rocky bench at the front of the shelter. **b**, Plan of Warraty rock shelter showing layout of excavation.

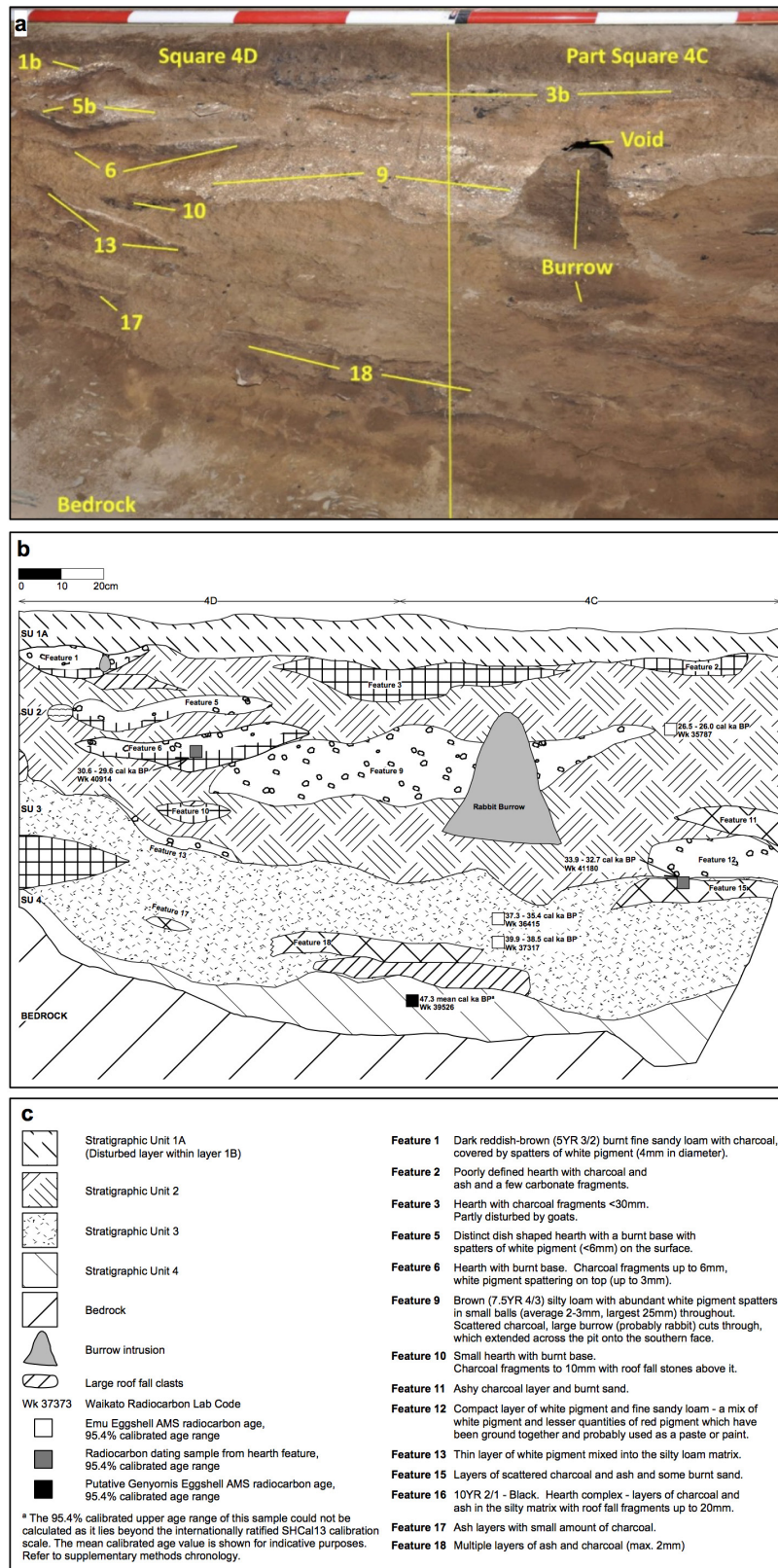
Excavation squares (4D/4C and 2C/2B) have a surface area of  $1 \text{ m}^2$ . Drawing by R. Frank. **c**, Warraty rock shelter excavation plan showing the layout of the excavation squares and  $25 \text{ cm} \times 25 \text{ cm}$  (A–D) quadrat units within each square.





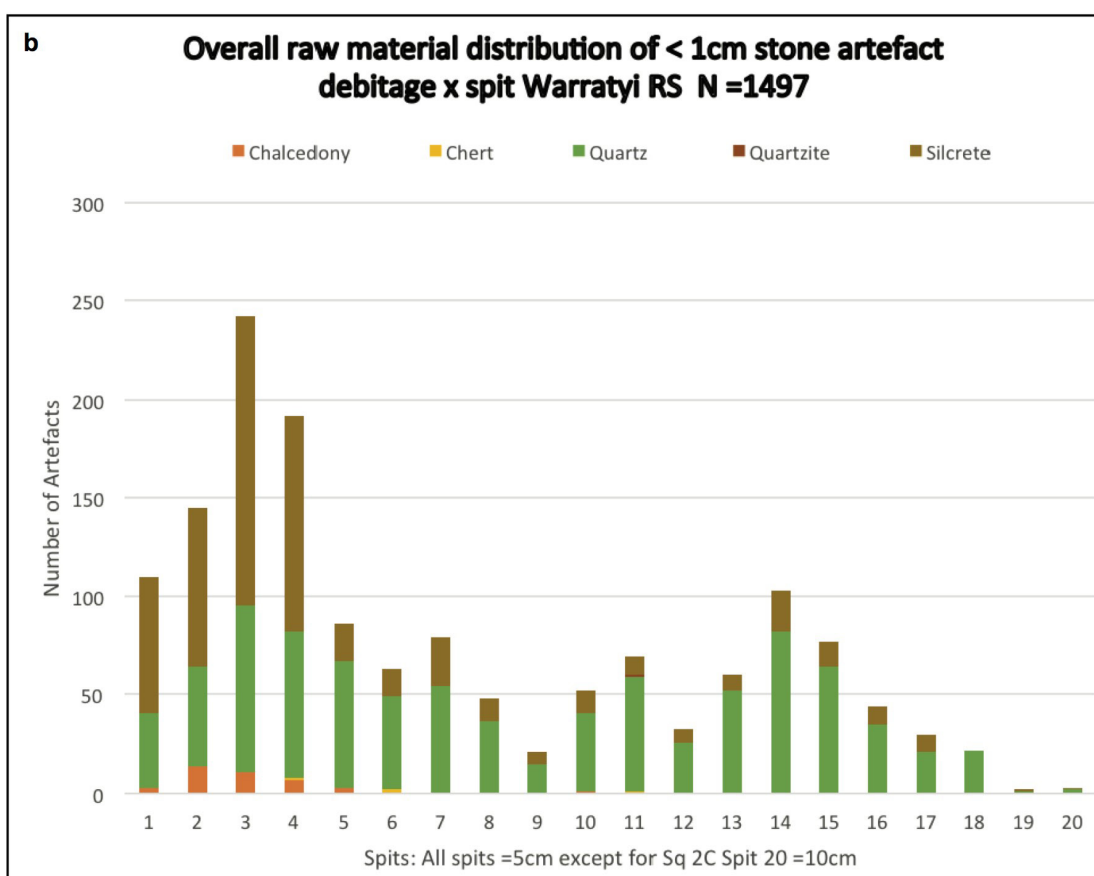
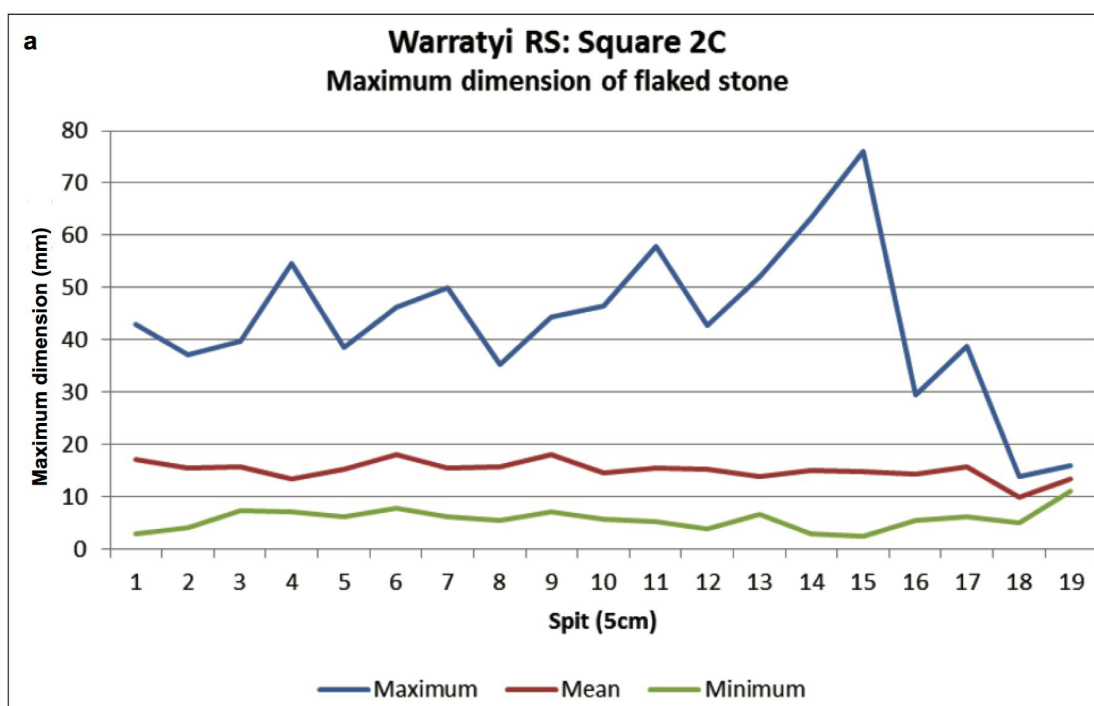
**Extended Data Figure 2 | Excavation profiles and section drawings of Warraty rock shelter square 2C.** **a**, Profile of the west face of square 2C from the surface to the weathered bedrock on the floor of the shelter. Several charcoal concentrations and small lenticular features of ash and charcoal were interpreted as hearths. Roof fall flakes occur throughout and the only evidence of burrowing is the small, darker grey area on the left edge, which was traced across the excavation and identified as a rabbit

burrow. This material was excavated separately to limit contamination of the main deposit. The numbers shown here denote the features shown in the profile drawing in **b**. The scale bar is marked with 20 cm units. **b**, Stratigraphic profile drawing of square 2C showing general stratigraphic features with feature descriptions and stratigraphic unit descriptions. **c**, Legend for stratigraphic profile drawing of square 2C. The scale bar is marked with 20 cm units.



**Extended Data Figure 3 | Excavation profiles and section drawings of Warraty rock shelter squares 4C and 4D.** **a**, Profile of the north face of square 4D and part of 4C (right of centre). All stratigraphic units are visible. A cross-section of the identified rabbit burrow is prominent in 4C; this disturbance, however, is limited and the burrow fill was excavated separately. Continuous concentrations of white pigment (9: gypsum spheres and pellets) are visible on either side of the burrow, attesting

to the stratigraphic integrity of the surrounding deposits. The numbers shown here correspond to the features noted in **b** and **c**, drawing and legend respectively. The scale bar is marked with 20 cm units. **b**, Profile drawing of excavated squares 4C and 4D, depicting the stratigraphy visible in **a**. A rabbit burrow is located in the centre of the drawing in square 4C. Vertical and horizontal scales are equal. **c**, Legend for stratigraphic profile drawing of squares 4C and 4D.

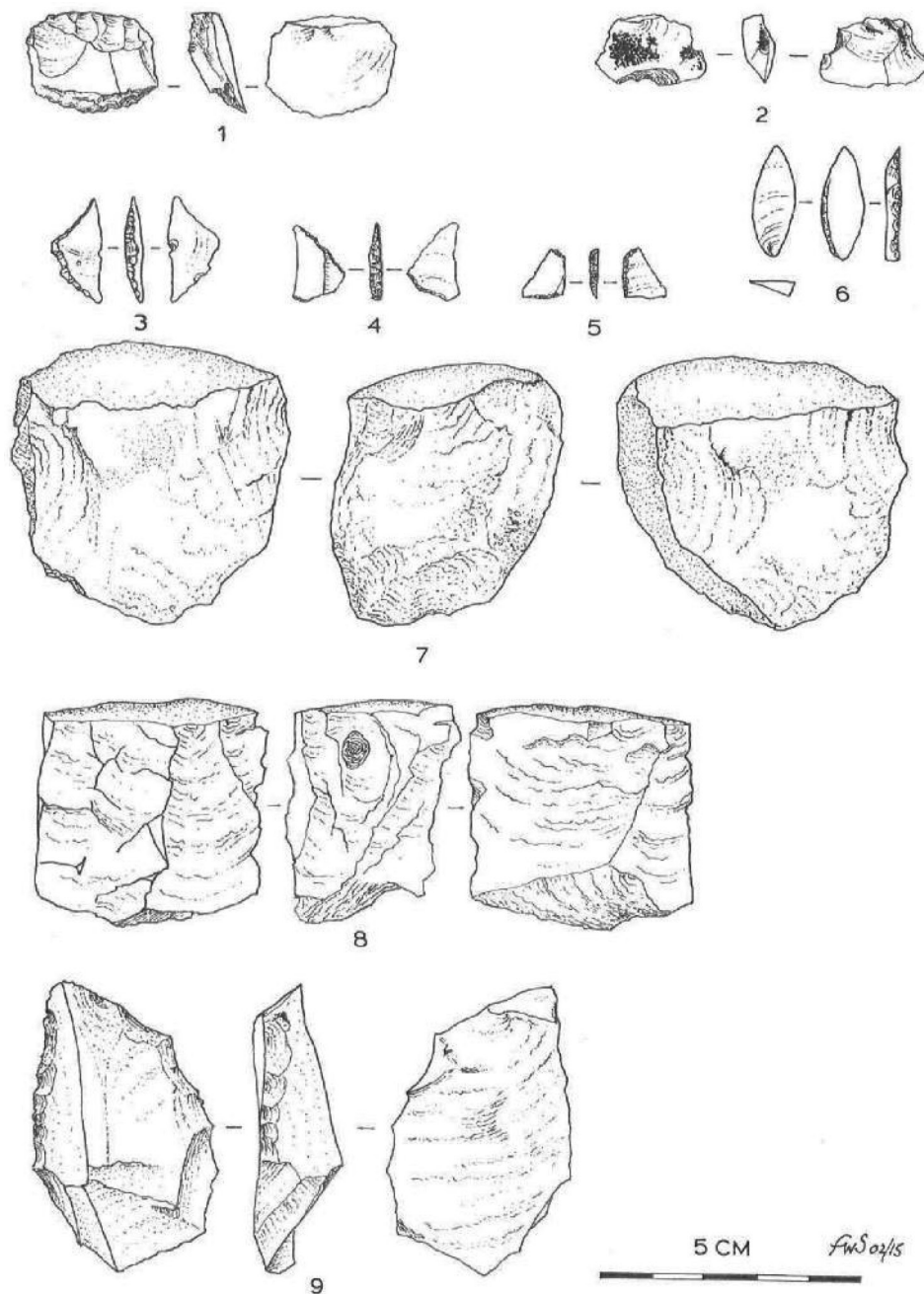


**Extended Data Figure 4 | Graphical representation of the size of stone artefacts and distribution of raw materials in Warratyi rock shelter.**

**a**, Distribution of lengths for *in situ* >1 cm artefacts found at Warratyi rock shelter. Blue line, maximum length; grey line, mean; orange, minimum

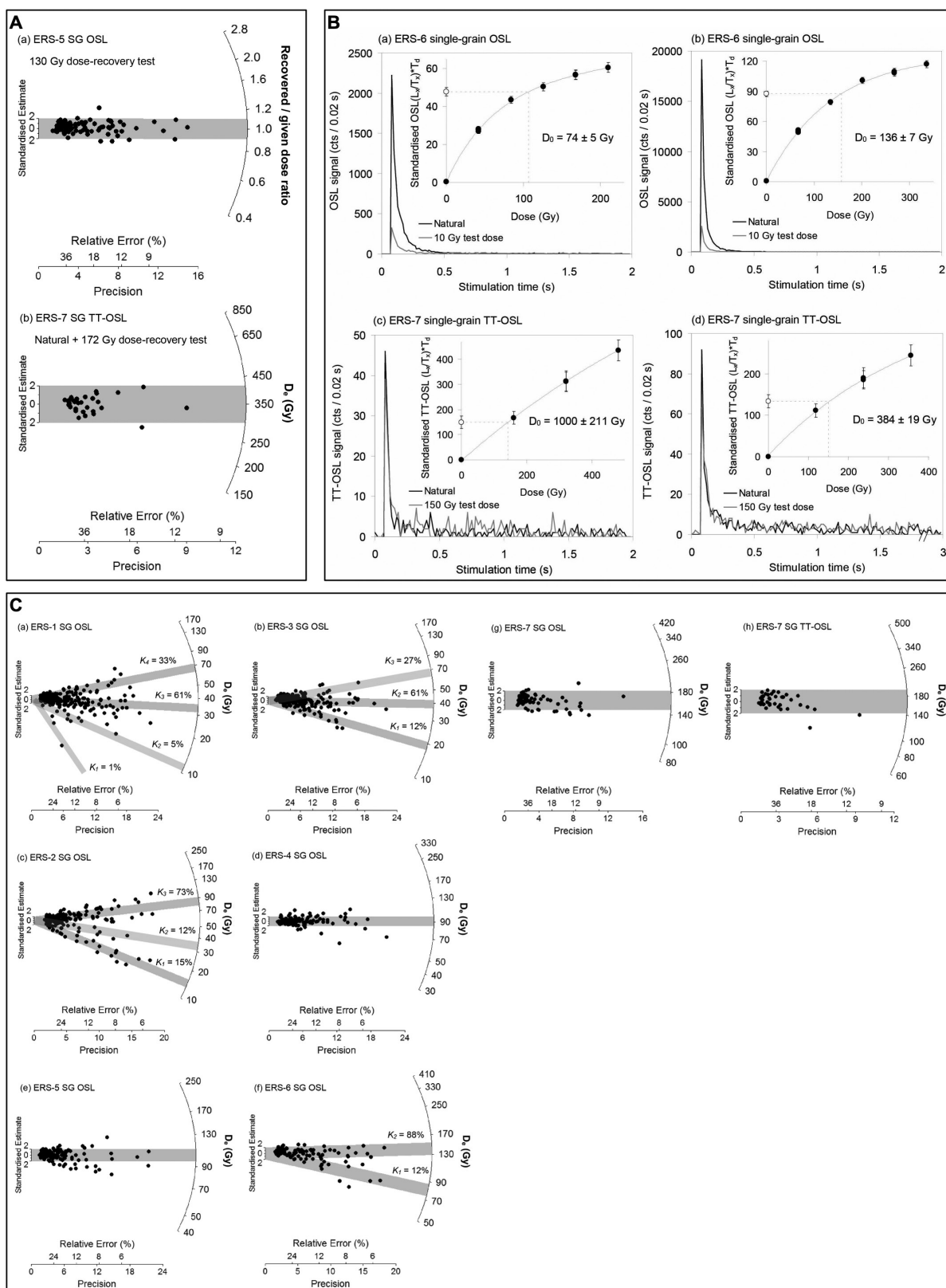
length. **b**, Distribution of raw material types for stone artefacts <1 cm in length by spit at Warratyi rock shelter. Orange, chalcedony; green, quartz; brown, silcrete; yellow, chert.





**Extended Data Figure 5 | A sample of the stone artefacts found in Warraty rock shelter.** Drawing of select Warraty stone artefacts. 1, Adze flake, chalcedony, square 2C, spit 4. 2, Adze flake, chalcedony, square 2C, spit 5, retains patches of resinous adhesive. 3, Geometric, backed artefact, silcrete, square 2B, spit 5. 4, Geometric, backed artefact, silcrete, square

2B, spit 5. 5, Geometric, backed artefact, silcrete, square 2B, spit 5. 6, Flake with backing on one edge, silcrete, square 2B, spit 5. 7, Core, rolled quartz pebble, square 4C, spit 9. 8, Core, rolled quartz pebble, square 4C, spit 15. 9, Retouched, quartzite flake, square 2C, spit 15.

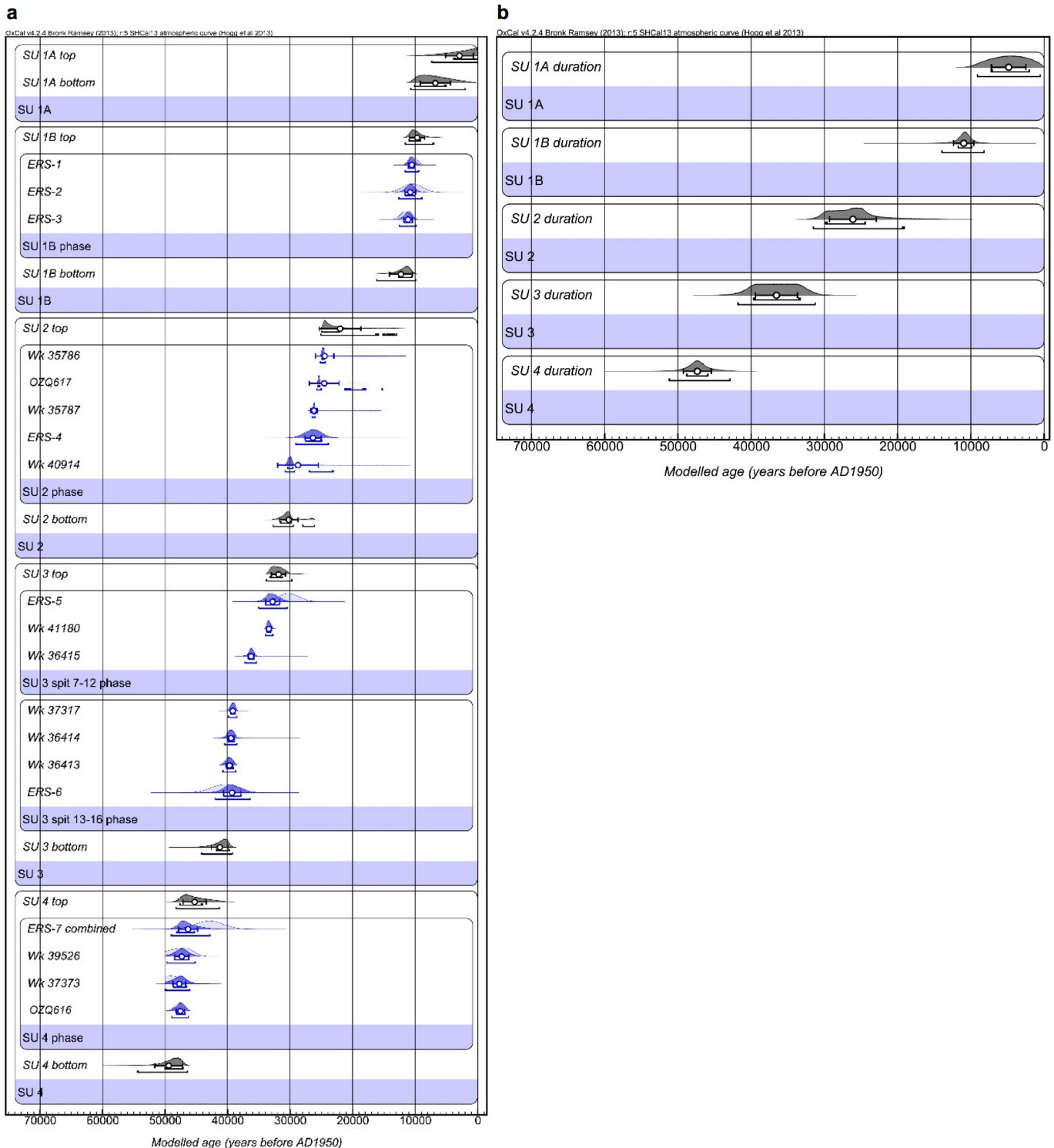


Extended Data Figure 6 | See next page for caption.

**Extended Data Figure 6 | Single-grain OSL dose-recovery and single-grain OSL and TT-OSL  $D_e$  distribution results for Warraty rock shelter** **A**, Single-grain OSL and TT-OSL dose-recovery test results. **Aa**, Radial plot showing the recovered to given dose ratios obtained for individual quartz grains of ESR-5 using the OSL SAR procedure (Supplementary Table 3). The natural OSL signals of these grains were first optically bleached with two 1,000-s blue LED illuminations at ambient temperature, each separated by a 10,000 s pause. The grey shaded region on the radial plot is centred on the administered dose for each grain (sample average = 130 Gy). **Ab**, Radial plot showing the dose-recovery test (natural + dosed)  $D_e$  values obtained for individual quartz grains of ESR-7 using the TT-OSL SAR procedure (Supplementary Table 3). The single-grain natural signals were not bleached during the TT-OSL dose-recovery test. Instead, a known dose of similar magnitude to the expected  $D_e$  was added on top of the natural signals. **B**, Representative OSL or TT-OSL decay and dose–response curves for individual quartz grains from Warraty rock shelter. **Ba**, Quartz grain from sample ERS-6 with an average

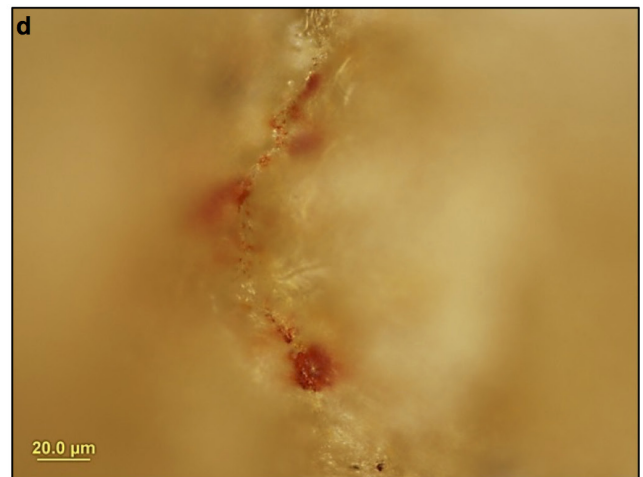
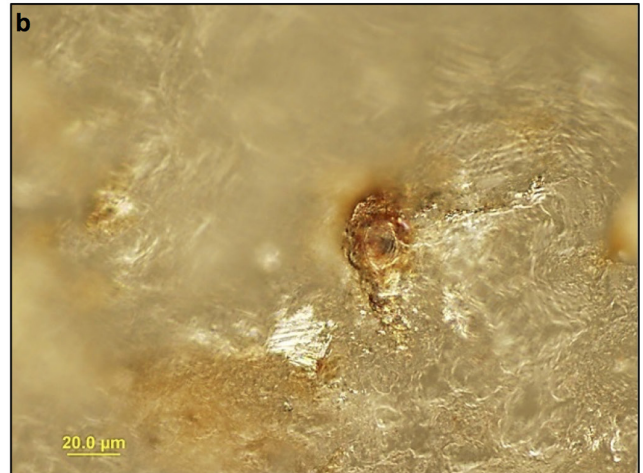
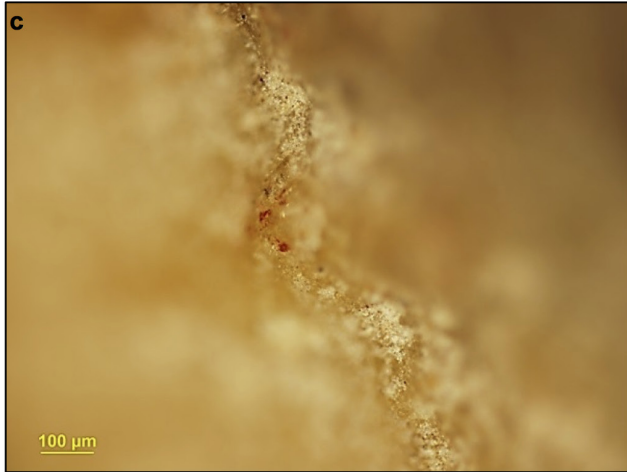
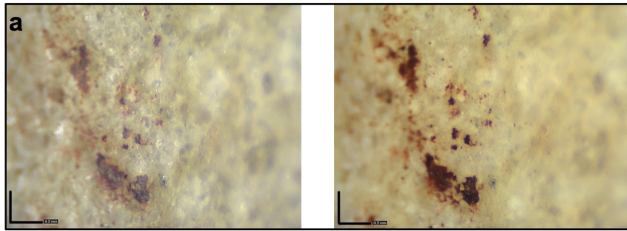
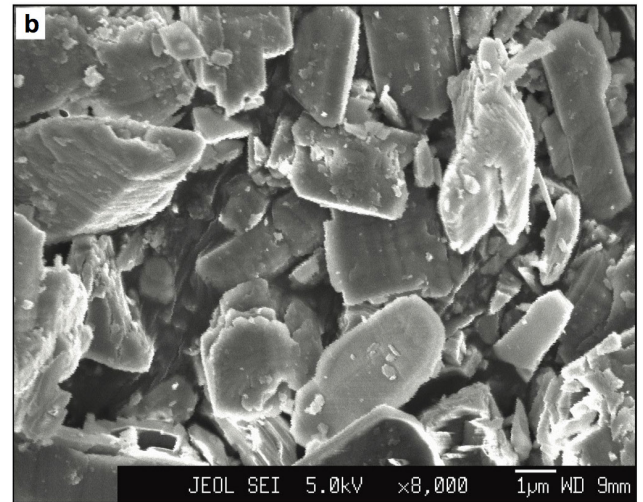
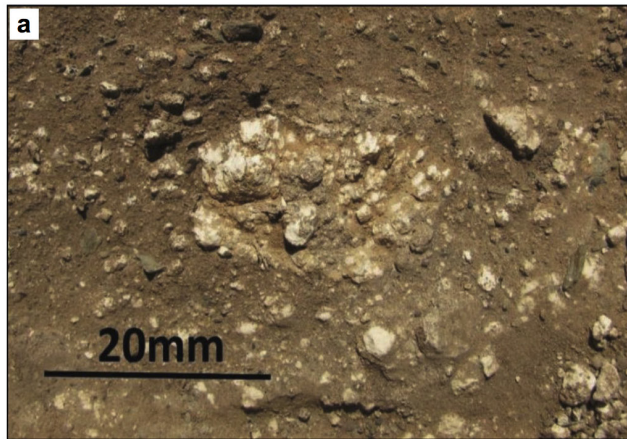
OSL signal ( $T_n \approx 800$  counts per 0.17 s). **Bb**, Quartz grain from sample ERS-6 with a relatively bright OSL signal ( $T_n \approx 5,000$  counts per 0.17 s). **Bc**, Quartz grain from sample ERS-7 with an average TT-OSL signal ( $T_n \approx 100$  counts per 0.17 s). **Bd**, Quartz grain from sample ERS-7 with a relatively bright TT-OSL signal ( $T_n \approx 200$  counts per 0.17 s). In the insets, open circles denote the sensitivity-corrected natural OSL or TT-OSL signals, and filled circles denote the sensitivity-corrected regenerated OSL or TT-OSL signals.  $D_0$  values characterize the rate of signal saturation with respect to administered dose and equate to the dose value for which the saturating exponential dose–response curve slope is  $1/e$  (or  $\sim 0.37$ ) of its initial value. **C**, Single-grain OSL and TT-OSL  $D_e$  distributions for the Warraty rock shelter samples. Grey bands are centred on the  $D_e$  values derived using either the central age model (samples ERS-4 (**Cc**), ERS-5 (**Ce**), ERS-7(**Cg**)), or the finite-mixture model (samples ERS-1 (**Ca**), ERS-2 (**Cb**), ERS-3 (**Cc**), ERS-6(**Cf**)). Percentage of grains associated with each finite-mixture model component ( $k_n$ ) are shown in **Ca**, **Cb**, **Cc** and **Cf**.





**Extended Data Figure 7 | Bayesian modelling results for Warraty rock shelter.** **a**, Bayesian modelling of optical-dating and  $^{14}\text{C}$  results from Warraty rock shelter. The prior age distributions for the dating samples (likelihoods) are shown as light blue probability density functions (PDFs). The modelled posterior distributions for the dating sample and stratigraphic unit boundaries are shown as dark blue and grey PDFs, respectively. Optical dating and  $^{14}\text{C}$  ages are shown on a calendar year timescale and both are expressed in years before 1950 AD. The  $^{14}\text{C}$  data were input into the model as conventional ages and were subsequently

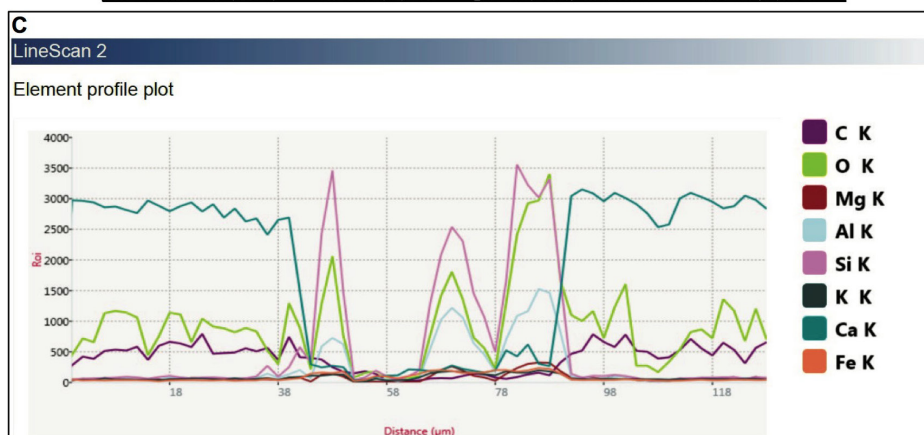
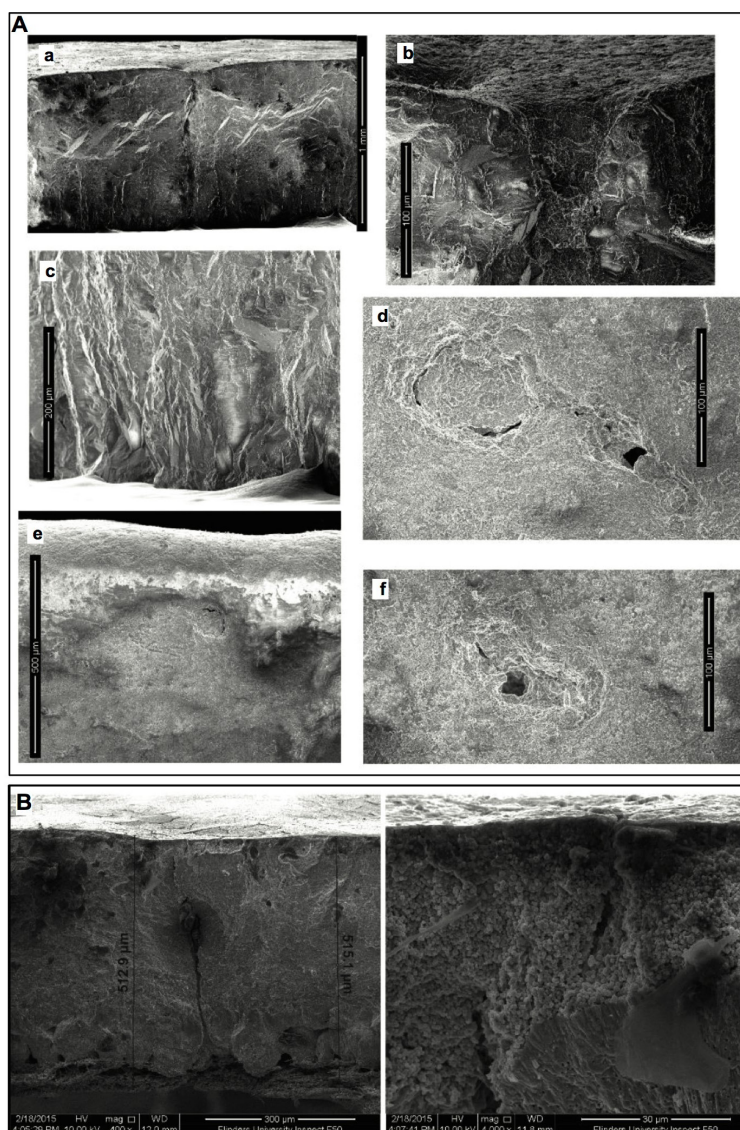
calibrated using the SHCal13 curve<sup>64</sup> as part of the modelling procedure. The white circles and associated error bars represent the mean ages and 1 standard error uncertainty ranges of the PDFs. The 68.2% and 95.4% ranges of the highest posterior probabilities are indicated by the horizontal bars underneath the PDFs. **b**, Bayesian-modelled durations of the stratigraphic units at Warraty rock shelter. The PDFs have been calculated from the modelled posterior probabilities of the upper and lower boundaries of each stratigraphic unit (shown as grey PDFs in plot **a** using the difference query function in OxCal v4.2.4).

**A****B**

**Extended Data Figure 8 | Distribution of resin and red ochre material on stone artefacts and gypsum analysis.** **Aa**, Resin adhering to the medial aris ridge that runs the length of the dorsal surface at 215× magnification in plane (pp, left) and cross polarized (xp, right). Scale bar, 0.2 mm. **Ab**, Resin with carbonized inclusions and associated hafting polish 500× darkfield. Scale bar, 20 μm. **Ac**, Ocre grains within the worked margin; 50× darkfield. Scale bar, 100 μm. **Ad**, Worked ochre grains; 500× brightfield cross polarized. Scale bar, 20.0 μm. **Ba**, Spheres of gypsum lying

*in situ* within excavated square 4C, Warraty rock shelter. **Bb**, Scanning electron microscope image of gypsum from Warraty rock shelter. The sample consisted of very fine-grained clear gypsum-cleavage plates, mostly sub-equidimensional to slightly tabular and 5–10 μm across. There are occasional larger tabular-acicular (needle-like) fragments up to 100 μm long. Scattered through the gypsum are well-rounded, sub-spherical silt-sized grains of quartz 50–75 μm across, generally reddish in colour. Scale bar, 1 μm.



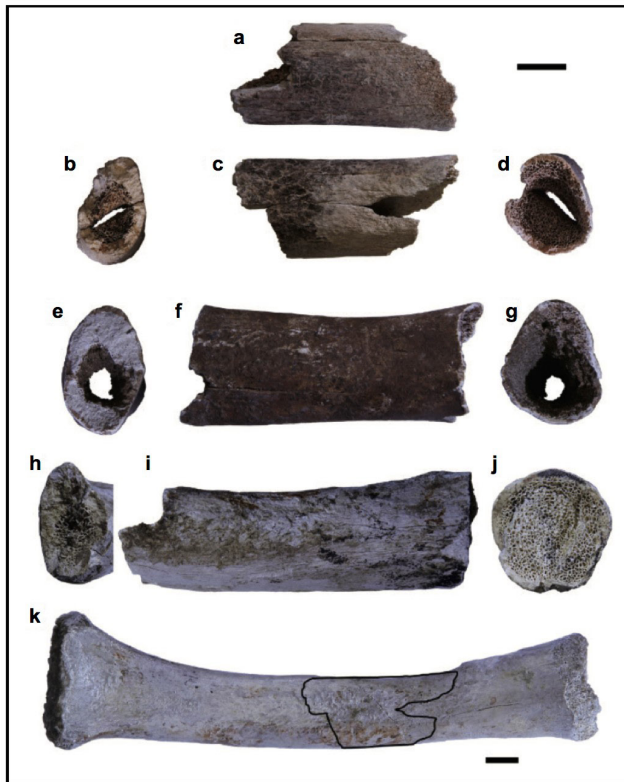


**Extended Data Figure 9 | Analysis of putative *Genyornis* oological eggshell material.** **A**, Scanning electron micrographs of sample 1 from square 4C, quadrat B, spit 18. **Aa**, view of a fresh broken edge showing a smooth eroded-out surface (upper), a section through a pore canal; **Ab**, detail of the pore canal opening showing the material filling pore; **Ac**, detail of the inner layer 1 and smooth inner surface of the shell showing that it is slightly eroded, so that the tips of the mammillary cones of layer 1 are lost; **Ad**, outer surface of the shell showing an elongated pair of pore canal openings occluded by material; **Ae**, outer surface of the shell showing the rounded eroded edge of the fragment; **Af**, outer surface of the shell and another opening to a pore canal showing its elongated nature

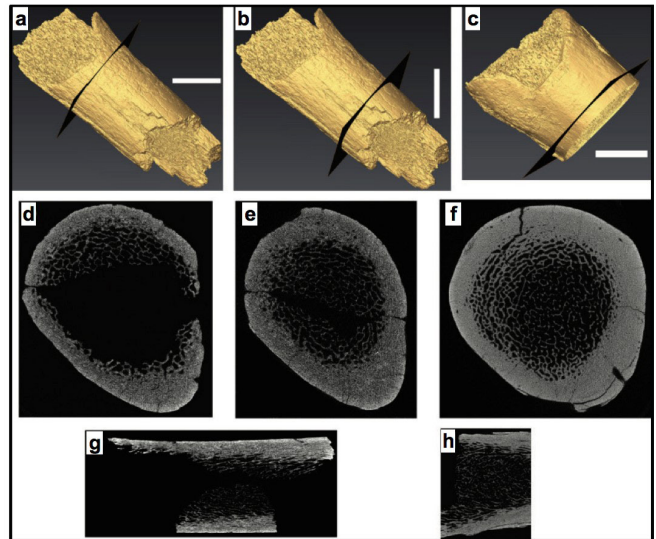
and occlusion by material. **B**, Scanning electron micrograph views of an eggshell of *Anseranas semipalmata* (sample from the South Australia Museum B.14591) showing the cross-section on the left and a detail of the accessory layer on the right. Note that the accessory layer comprises an amorphous mass of similar sized spheres, a structure that typifies the accessory layer of many galloanseres, including that of putative *Genyornis* material. **C**, Element-profile plot across the pore aperture shown in **Ab**. The sides of the pore are at approximately 40 and 90  $\mu\text{m}$  on the  $x$  axis and on either side the profile reflects the dominant  $\text{CaCO}_3$  nature of eggshell. Within the pore, elevated levels of iron (Fe), silica (Si) and aluminium (Al) are present.



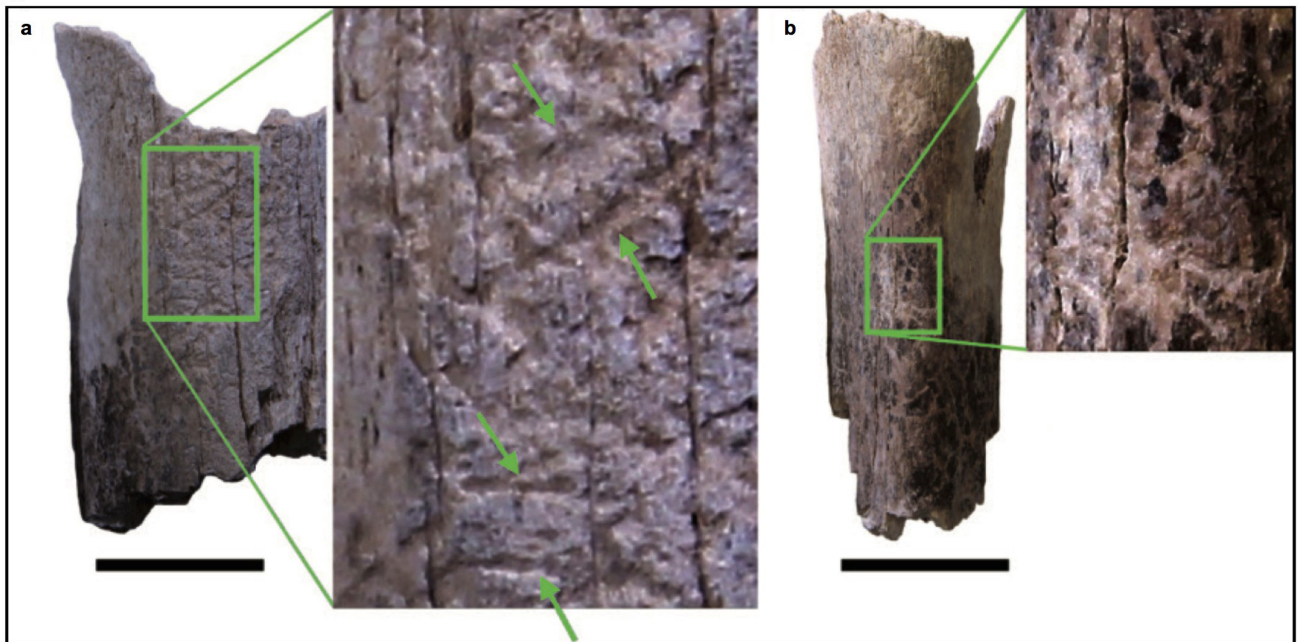
A



B



C



#### Extended Data Figure 10 | Megafaunal bone evidence of *D. optatum*.

**A**, Comparison between the *Diprotodon* bone (EMU RS-6737-7754) from Warraty rock shelter and juvenile radius specimens of *D. optatum* from the South Australian Museum (SAMA P51340–P51342). **Aa**, EMU RS-6737-7754 in anterior view. **Ab**, View of the broken distal end of EMU RS-6737-7754. **Ac**, EMU RS-6737-7754 in posterior view. **Ad**, View of the broken proximal end of EMU RS-6737-7754. **Ae**, View of the broken distal end of SAMA P51340. **Af**, SAMA P51340 in posterior view. **Ag**, View of the broken proximal end of SAMA P51341. **Ah**, View of the broken distal end of SAMA P51341. **Ai**, SAMA P51341 in posterior view. **Aj**, View of the broken proximal end of SAMA P51342. **Ak**, SAMA P51342 in posterior view. The silhouette represents the approximate corresponding position for EMU RS-6737-7754. Scale bars, 20 mm. The top scale bar is

for panels **Aa–Aj**; the bottom scale bar is for panel **Ak**. **B**, Micro-computed tomographic images comparing EMU RS-6737-7754 with a juvenile radius specimen of *D. optatum*. **Ba**, Isosurface rendering of EMU RS-6737-7754 showing the position of a more proximal cross-section. **Bb**, Isosurface rendering of EMU RS-6737-7754 showing the position of a more distal cross-section. **Bc**, Isosurface rendering of SAMA P51341 showing the position of the cross-section. **Bd**, More proximal cross-section of EMU RS-6737-7754 (as shown in **Ba**). **Be**, More distal cross-section of EMU RS-6737-7754 (as shown in **Bb**). **Bf**, Cross-section of SAMA P51341 (as shown in **Bc**). **Bg**, Longitudinal-section of EMU RS-6737-7754. **Bh**, Longitudinal-section of a portion of SAMA P51341. Scale bars, 20 mm. **C**, Taphonomic markings on the surface of EMU RS-6737-7754.

# A brain–spine interface alleviating gait deficits after spinal cord injury in primates

Marco Capogrosso<sup>1,2\*</sup>, Tomislav Milekovic<sup>1\*</sup>, David Borton<sup>1,3\*</sup>, Fabien Wagner<sup>1</sup>, Eduardo Martin Moraud<sup>2</sup>, Jean-Baptiste Mignardot<sup>1</sup>, Nicolas Buse<sup>4</sup>, Jerome Gandar<sup>1</sup>, Quentin Barraud<sup>1</sup>, David Xing<sup>3</sup>, Elodie Rey<sup>1</sup>, Simone Duis<sup>1</sup>, Yang Jianzhong<sup>5</sup>, Wai Kin D. Ko<sup>5</sup>, Qin Li<sup>5,6</sup>, Peter Detemple<sup>7</sup>, Tim Denison<sup>4</sup>, Silvestro Micera<sup>2,8</sup>, Erwan Bezard<sup>5,6,9,10</sup>, Jocelyne Bloch<sup>11</sup> & Grégoire Courtine<sup>1,11</sup>

Spinal cord injury disrupts the communication between the brain and the spinal circuits that orchestrate movement. To bypass the lesion, brain–computer interfaces<sup>1–3</sup> have directly linked cortical activity to electrical stimulation of muscles, and have thus restored grasping abilities after hand paralysis<sup>1,4</sup>. Theoretically, this strategy could also restore control over leg muscle activity for walking<sup>5</sup>. However, replicating the complex sequence of individual muscle activation patterns underlying natural and adaptive locomotor movements poses formidable conceptual and technological challenges<sup>6,7</sup>. Recently, it was shown in rats that epidural electrical stimulation of the lumbar spinal cord can reproduce the natural activation of synergistic muscle groups producing locomotion<sup>8–10</sup>. Here we interface leg motor cortex activity with epidural electrical stimulation protocols to establish a brain–spine interface that alleviated gait deficits after a spinal cord injury in non-human primates. Rhesus monkeys (*Macaca mulatta*) were implanted with an intracortical microelectrode array in the leg area of the motor cortex and with a spinal cord stimulation system composed of a spatially selective epidural implant and a pulse generator with real-time triggering capabilities. We designed and implemented wireless control systems that linked online neural decoding of extension and flexion motor states with stimulation protocols promoting these movements. These systems allowed the monkeys to behave freely without any restrictions or constraining tethered electronics. After validation of the brain–spine interface in intact (uninjured) monkeys, we performed a unilateral corticospinal tract lesion at the thoracic level. As early as six days post-injury and without prior training of the monkeys, the brain–spine interface restored weight-bearing locomotion of the paralysed leg on a treadmill and overground. The implantable components integrated in the brain–spine interface have all been approved for investigational applications in similar human research, suggesting a practical translational pathway for proof-of-concept studies in people with spinal cord injury.

A century of research in spinal cord physiology has demonstrated that circuits embedded in the lumbar segments of mammals can produce coordinated patterns of leg motor activity without brain input<sup>11,12</sup>. Various neuromodulation approaches have been developed to activate these circuits after injury to re-establish locomotion<sup>8,13–17</sup>. For example, epidural electrical stimulation (EES) of lumbar segments restored adaptive locomotion in paralysed rats<sup>8</sup>. Recent studies showed that EES is also capable of activating lumbar spinal circuits in people with paraplegia<sup>14,16</sup>.

These empirical observations prompted us to develop an evidenced-based framework to understand the interactions between EES and

spinal circuits<sup>8–10</sup>. We aimed to exploit this knowledge to optimize stimulation protocols for clinical applications. Computational modelling and functional experiments revealed that EES engages spinal circuits through the modulation of proprioceptive feedback circuits<sup>10</sup>. This framework guided the design of spatiotemporal neuromodulation therapies that not only activate but also control the activity of spinal circuits engaging synergistic muscle groups<sup>8–10</sup>, enabling robust modulation of locomotor movements in rats whose spinal cords were void of brain input.

However, volitional locomotion requires the brain to control the activity of spinal circuits. Brain–computer interface technologies<sup>1–4,18</sup> provide the tools to link the intended motor states to EES protocols<sup>19–21</sup> to re-establish voluntary control of locomotion after injury. For these developments, nonhuman primates are more appropriate models than rodents because they exhibit cortical engagement during locomotion similar to humans<sup>22</sup>, analogous recovery mechanisms from injury<sup>23</sup>, and comparable technological requirements<sup>24</sup>. Here, we decoded motor states from leg motor cortex activity to trigger EES protocols facilitating extension and flexion of the corresponding leg. We show that this brain–spine interface alleviated gait deficits after spinal cord injury in nonhuman primates.

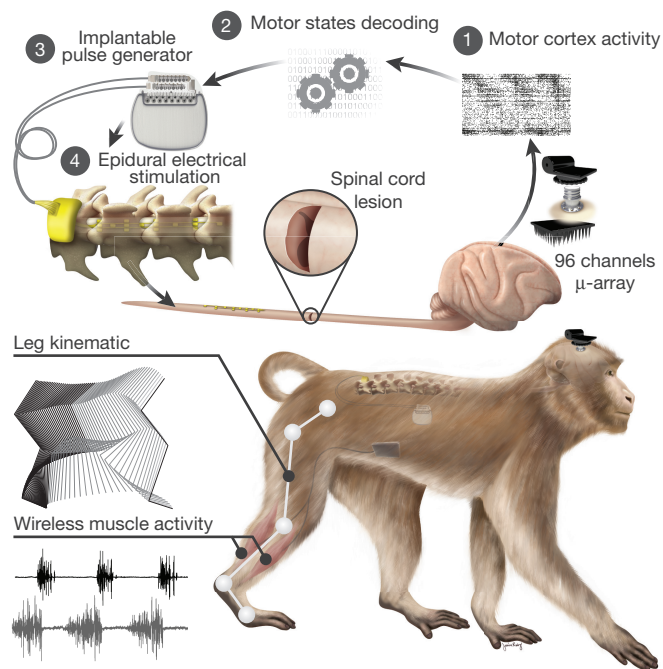
To support the development of the brain–spine interface, we established a wireless recording and stimulation platform in freely behaving, unconstrained and untethered nonhuman primates (Fig. 1 and Supplementary Video 1). Rhesus monkeys (Supplementary Table 1) were implanted with a microelectrode array into the leg area of the left motor cortex to record spiking activity from neuronal ensembles. Electromyographic signals were monitored using bipolar electrodes implanted into antagonist muscles spanning each joint of the right leg. Wireless modules enabled transmission of neural (20 kHz) and electromyographic (2 kHz) signals to external receivers<sup>25</sup>. We simultaneously acquired video recordings (100 Hz) to reconstruct whole-body kinematics<sup>23</sup>. To deliver EES, we used technologies previously developed in rats<sup>9</sup>, which we adapted to the characteristics of spinal segments and vertebrae measured in three monkeys (Extended Data Fig. 1). These spinal implants were inserted into the epidural space over lumbar segments, and connected to an implantable pulse generator commonly used for deep brain stimulation therapy. We engineered wireless communication modules that enabled control over the spatial and temporal parameters of EES with a latency of about 100 ms (Extended Data Fig. 2).

We first used well established methods<sup>9,26</sup> to identify the natural spatiotemporal pattern of motoneuron activation underlying locomotion. Our aim was to reproduce this pattern after injury.

<sup>1</sup>Center for Neuroprosthetics and Brain Mind Institute, School of Life Sciences, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. <sup>2</sup>Center for Neuroprosthetics and Institute of Bioengineering, School of Bioengineering, EPFL, Lausanne, Switzerland. <sup>3</sup>School of Engineering, Brown University, Providence, Rhode Island, USA. <sup>4</sup>Medtronic, Minneapolis, Minnesota, USA. <sup>5</sup>Motac Neuroscience Ltd, Manchester, UK. <sup>6</sup>Institute of Lab Animal Sciences, China Academy of Medical Sciences, Beijing, China. <sup>7</sup>Mainz Institute for Microtechnology, Fraunhofer Institute for Chemical Technology (ICT-IMM), Mainz, Germany. <sup>8</sup>The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy. <sup>9</sup>Institut des Maladies Neurodégénératives, University of Bordeaux, UMR 5293, Bordeaux, France. <sup>10</sup>CNRS, Institut des Maladies Neurodégénératives, UMR 5293, Bordeaux, France. <sup>11</sup>Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland.

\*These authors contributed equally to this work.

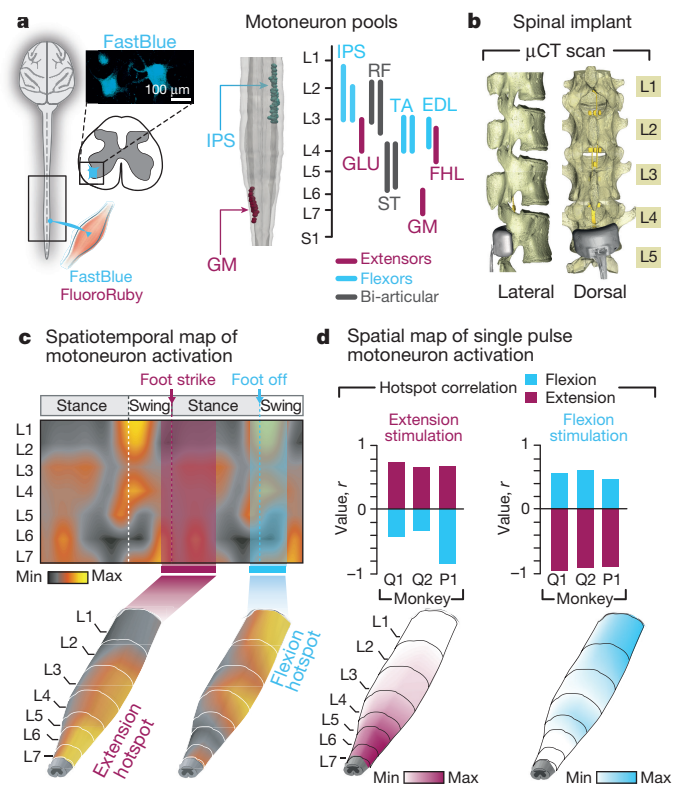




**Figure 1 | Conceptual and technological design of the brain–spine interface.** The monkeys were implanted with a microelectrode array into the leg area of the left motor cortex. During recordings, a wireless module transmitted broadband neural signals to a control computer. (1) A raster plot recorded over three successive gait cycles. Each line represents spiking events identified from one electrode, while the horizontal axis indicates time. (2) A decoder running on the control computer identified motor states from these neural signals. (3) These motor states triggered electrical spinal cord stimulation protocols. For this, the monkeys were implanted with a pulse generator featuring real-time triggering capabilities. (4) The stimulator was connected to a spinal implant targeting specific dorsal roots of the lumbar spinal cord. Electromyographic signals of an extensor (grey) and flexor (black) muscles acting at the ankle recorded over three successive gait cycles are shown together with a stick diagram decomposition of leg movements during the stance (grey) and swing (black) phases of gait. Copyright Jemère Ruby (2016).

We conducted an anatomical tracing to identify the spatial distribution of motoneuron pools innervating antagonist muscles spanning each joint of the leg (Fig. 2a). We then projected the muscle activity recorded during locomotion onto motoneuron locations to visualize the spatiotemporal maps of motoneuron activation (Fig. 2c). These maps showed that locomotion involves the successive activation of well defined hotspots located in specific regions of the spinal cord that were similar across monkeys (Extended Data Fig. 3). The most intense hotspots emerged in the caudal (L6/L7) and rostral (L1/L2) compartments of lumbar segments around the transitions between stance and swing phases. We labelled these hotspots the extension and flexion hotspots, respectively.

EES activates motoneurons through the recruitment of large-diameter proprioceptive fibres within the dorsal roots<sup>10,27</sup>. To access the extension and flexion hotspots, we targeted the dorsal roots projecting to spinal segments containing these hotspots. We reconstructed the spatial trajectory of the dorsal roots innervating each lumbar segment, and integrated this information together with motoneuron distribution into a unified library (Fig. 2a). We used the entry points of the dorsal roots as the targeted anatomical landmarks that guided the design and positioning of spinal implants (Fig. 2b and Extended Data Fig. 1). Experiments in three sedated monkeys confirmed that single EES pulses delivered through the electrodes targeting the extension and flexion hotspots led to spinal segment activation that correlated with the activation of these hotspots during locomotion (Fig. 2c and d, and Extended Data Fig. 3).

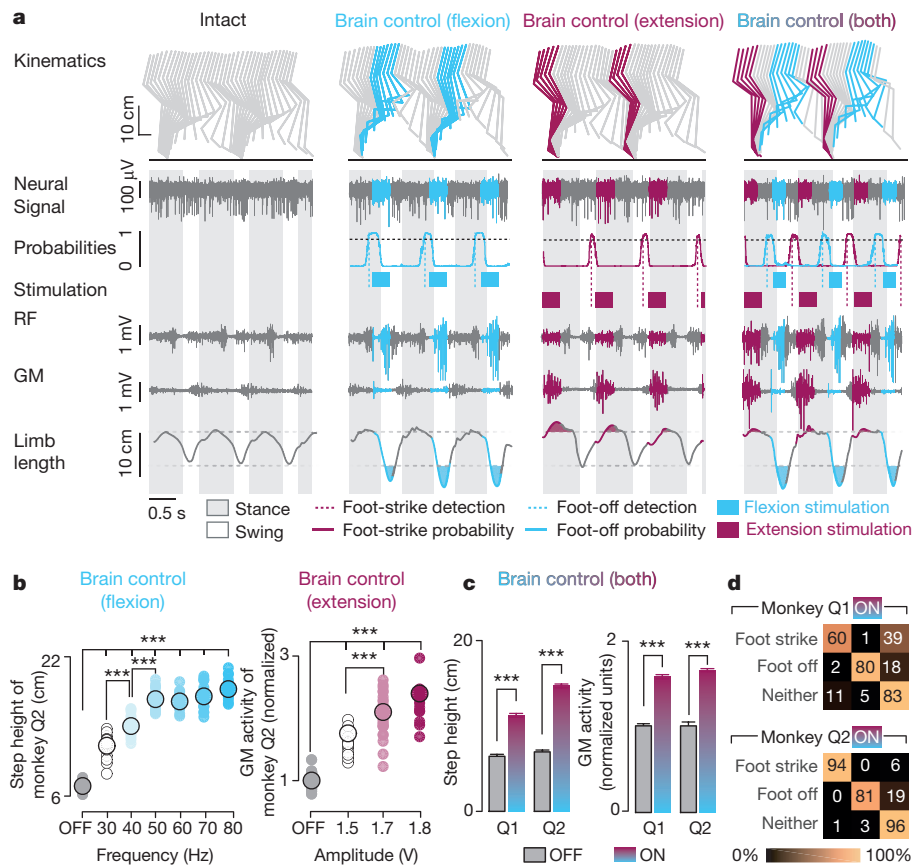


**Figure 2 | Development and validation of selective spinal cord stimulation protocols.** **a**, Diagram illustrating injections of retrograde anatomical tracers into leg muscles to label motoneurons. The inset shows a confocal photograph of labelled motoneurons. The 3D reconstruction shows each labelled motoneuron innervating the iliopsoas (IPS) and gastrocnemius medialis (GM) muscles. The same procedure was applied to the gluteus medius (GLU), rectus femoris (RF), semitendinosus (ST), extensor digitorum longus (EDL), and flexor hallucis longus (FHL) muscles. The diagram reports the distribution of leg motoneurons within the spinal cord ( $n = 1–2$  monkeys per muscle). **b**, Representative micro-computed tomography ( $\mu$ CT) scans of the spinal implant (monkey Q1). **c**, Electromyography of the recorded leg muscles (monkey Q1) was projected onto the motoneuron locations in the spinal cord to compute the mean ( $n = 73$  gait cycles) spatiotemporal map of motoneuron activation during locomotion. Maps recorded around foot off ( $-10\%$  to  $+20\%$  of gait cycle) and foot strike ( $-10\%$  to  $+30\%$ ) were extracted to highlight extension and flexion hotspots (see data from other monkeys in Extended Data Fig. 3). **d**, Median ( $n = 6$  pulses) spatial map of motoneuron activation resulting from single pulses of stimulation delivered through the electrodes targeting the extension and flexion hotspots. For each monkey, the bar plots report the correlation between these spatial maps and the maps corresponding to the extension and flexion hotspots. L, lumbar.

We next exploited cortical signals to decode the temporal structure of extensor and flexor hotspot activation. The spiking activity recorded from the left motor cortex displayed cyclic modulations that were phase-locked with right leg movements (Extended Data Fig. 4a). We developed a decoder that calculated the probability of foot-strike and foot-off events from this modulation to anticipate the activation of extensor and flexor hotspots associated with right leg movements (Extended Data Fig. 4b). Evaluations in two intact monkeys showed that the decoder accurately predicted these gait events in real time over extended periods of locomotion, including when initiating and terminating gait, and during rest (Extended Data Fig. 5).

We then exploited our wireless platform to implement a brain–spine interface—a system wherein the decoded motor states triggered EES protocols targeting the extension and flexion hotspots. We tested the capacity of the brain–spine interface to modulate the extension and flexion hotspots independently and simultaneously in two intact monkeys during locomotion on a treadmill. We calibrated the decoder with





**Figure 3 | Brain-controlled stimulation modulates the extension and flexion of the leg during locomotion in intact monkeys.** **a**, Two successive gait cycles recorded during locomotion without stimulation and during brain-controlled stimulation of the flexion hotspot, extension hotspot, or both (monkey Q2). From top to bottom: stick diagram decompositions of right leg movements; example of single-channel neural recording; probability of foot-off and foot-strike motor states; detected motor states (cyan and magenta broken lines), periods of stimulation through the electrodes targeting the flexion and extension hotspots; electromyographic signals; and limb length calculated as the distance from the hip to the fifth metatarsal joint. The grey and white backgrounds correspond to stance and swing, respectively. **b**, Relationship between

temporal offsets that were tuned to trigger and terminate stimulation protocols concomitantly to the activation of each hotspot (Fig. 3a and Supplementary Methods). We used data without and with stimulation to calibrate the decoders<sup>4</sup>, which substantially improved decoding accuracy (Extended Data Fig. 5).

Without prior training of the monkeys, brain-controlled stimulation of the extension and flexion hotspots immediately modulated kinematic and muscle activity parameters related to the extension and flexion of the leg ipsilateral to stimulation (Fig. 3). A gradual increase in the frequency or amplitude of EES pulses led to a monotonic modulation of these parameters (Extended Data Fig. 6). We previously documented similar responses in rodents<sup>8–10</sup>, suggesting that the mechanisms underlying the modulation of spinal activity with EES are similar across mammals, including humans<sup>14,16</sup>.

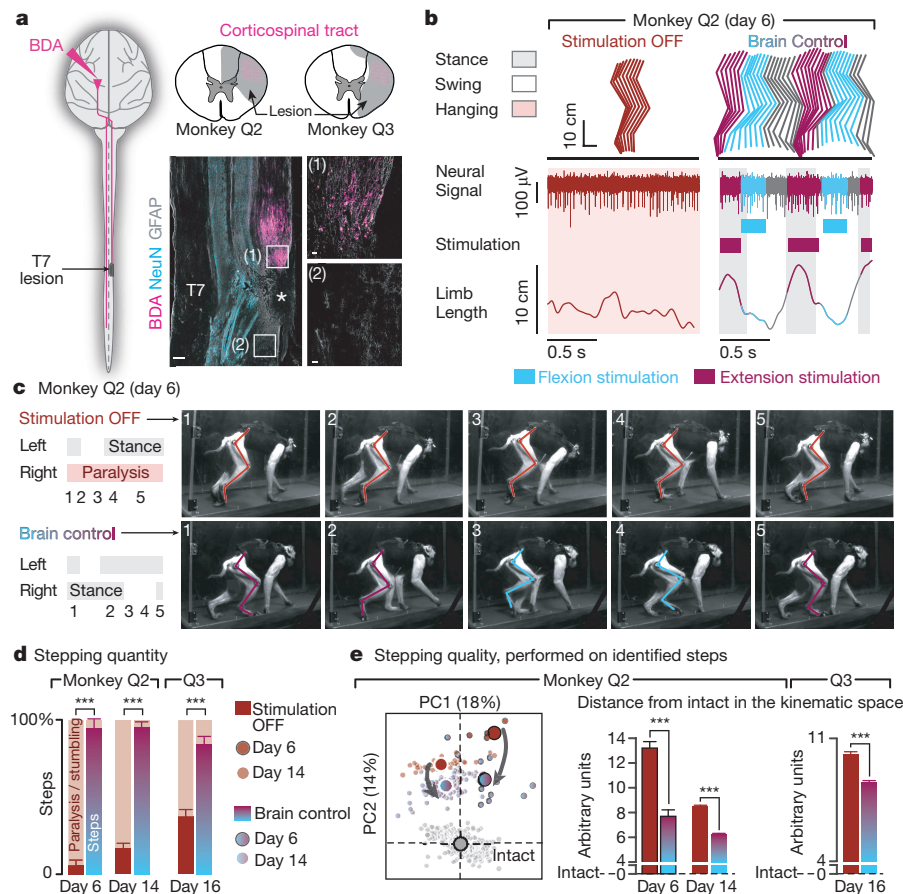
Finally, we tested the ability of the brain–spine interface to alleviate locomotor deficits after a lesion of the corticospinal tract extending into the right dorsolateral column of mid-thoracic segments in two monkeys (Fig. 4a). Additional pathways were damaged, including the rubrospinal tract, dorsal column and reticulospinal fibres. This lesion initially led to a paralysis of the leg ipsilateral to the lesion, followed by an extensive yet incomplete spontaneous recovery (Fig. 4b and Extended Data Fig. 7). During the first week after the lesion and without training of the monkeys, the brain–spine interface restored weight-bearing locomotion

frequency of brain-controlled flexion stimulation and step height, and between the stimulation amplitude and the activity of the GM muscle during brain-controlled extension stimulation (normalized to the GM mean activity 1 for stimulation ‘OFF’). Values for individual gait cycles (smaller dots) and mean values (larger dots) are shown. \*\*\* $P < 0.001$ , Wilcoxon rank-sum test. **c**, Bar plots reporting the mean step height and mean GM activity without stimulation and during brain-controlled stimulation for monkeys Q1 ( $n = 125$  steps) and Q2 ( $n = 119$  steps). \*\*\* $P < 0.001$ , Wilcoxon rank-sum test. Error bars, s.e.m. **d**, Decoder confusion matrices calculated during brain-controlled stimulation for monkeys Q1 ( $n = 125$  steps) and Q2 ( $n = 119$  steps).

on a treadmill (Fig. 4b, c) and overground (Extended Data Fig. 8), improving both the quantity and quality of steps performed by the impaired leg (Fig. 4d, e and Extended Data Fig. 9 and Supplementary Video 1). The quantity and quality of steps was directly linked to the temporal structure of the stimulation (Extended Data Fig. 10).

Accuracy of decoded extension and flexion motor states from neural activity declined shortly after the lesion and then improved during the following week. This improvement suggested that the initial decrease was probably due to the reorganization of cortical dynamics (Extended Data Fig. 9). This recovery coincided with improvement in the quantity and quality of steps, indicating that the monkeys had spontaneously regained some degree of neural control over the impaired leg (Extended Data Fig. 7). At this stage, the brain–spine interface alleviated many of the remaining gait deficits (Fig. 4d, e). Tuning the EES frequency maximized the quantity and quality of steps, whereas the same stimulation protocols (that is, the same electrode configurations at same amplitudes and frequencies of stimulation) applied continuously failed to facilitate locomotion or were markedly less efficient than brain-controlled stimulation (Extended Data Fig. 10).

The recovery of coordinated, weight-bearing locomotion in a primate model of spinal cord injury emphasizes the therapeutic potential of the brain–spine interface for clinical applications. We have integrated intracortical arrays<sup>2,3</sup>, wireless modules<sup>25</sup> and pulse generators that have



**Figure 4 | The brain–spine interface alleviates gait deficits after spinal cord injury.** **a**, Diagram illustrating the location of the lesion and corticospinal tract labelling using biotinylated dextran amine (BDA). Right, anatomical reconstructions of spinal segments containing the lesion (grey), for monkeys Q2 and Q3. Photographs including insets (1) and (2), showing a longitudinal view of the lesioned spinal cord wherein astrocytes (glial fibrillary acidic protein (GFAP); grey), neurons (anti-NeuN antibody (NeuN); cyan) and corticospinal tract axons (BDA, pink) are labelled. The asterisk indicates the lesion. The overview scale bar is 500  $\mu$ m and the inset scale bars are 50  $\mu$ m. **b**, Gait cycles performed during locomotion without stimulation and during brain-controlled stimulation of both flexion and extension hotspots in monkey Q2 at 6 days post-lesion. Conventions are as in Fig. 3. Limb paralysis is shown in red. **c**, Snapshots extracted from video recordings showing a sequence of leg movements without stimulation and

during brain-controlled stimulation (monkey Q2, 6 days post-injury). The timeline indicates video snapshot timing. The key refers to panels **d** and **e**. **d**, Bar plots reporting the ratio between steps performed by the affected versus unaffected leg by each monkey without stimulation ( $n = 6$  for day 6 and  $n = 39$  for day 14 for Q2,  $n = 68$  for Q3) and during brain-controlled stimulation ( $n = 12$  for day 6 and  $n = 93$  for day 14 for Q2,  $n = 31$  for Q3). \*\*\* $P < 0.001$ , bootstrap. Error bars, s.e.m. **e**, Principal component analysis applied on 26 gait parameters for Q2. All the gait cycles corresponding to limb paralysis or stumbling have been excluded from this analysis. Each gait cycle is shown in the space defined by PC1 and PC2. Bar plots reporting the mean Euclidean distance between pre-lesion and post-lesion gait cycles corresponding to steps, calculated in the entire kinematic space. \*\*\* $P < 0.001$ , Wilcoxon rank-sum test. Error bars, s.e.m.

been approved for research applications in humans, thus potentially enabling proof-of-concept clinical studies.

Our brain–spine interface exploits neuronal ensemble modulation that naturally occurs during locomotion, immediately linking cortical dynamics with spatiotemporal neuromodulation therapies without prior training of the monkeys. This ecological approach<sup>28</sup> enabled a smooth cooperation between residual supraspinal signals and the brain–spine interface in generating leg movements. Imaging<sup>29</sup> and electrophysiological<sup>4</sup> studies have shown that leg motor cortex dynamics is preserved in people with paralysis. Moreover, cortical activity modulates with intended movements in people with long-lasting tetraplegia, which has allowed them to control robotic arms<sup>2,3</sup> and neuromuscular stimulators<sup>4</sup>. These results suggest that the decoding strategy employed in this study may have useful application in people with paraplegia.

Our model of paralysis avoided many of the complications associated with severe injuries that are difficult to manage and ethically debatable in primates<sup>24</sup>. The use of a brain–spine interface to restore bipedal locomotion in humans after severe injuries may require additional interventions, including monoaminergic replacement therapies<sup>13,19</sup> compensating for the interrupted source of serotonin from brainstem

centres and robotic systems to sustain balance. Nevertheless, several people with complete leg paralysis regained weight-bearing standing and stepping movements during continuous EES<sup>14,16</sup>. Therefore, it should now be possible to test the efficacy of the brain–spine interface to enhance neuroplasticity<sup>21,30</sup> during rehabilitation in people with spinal cord injury.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 10 May; accepted 27 September 2016.**

- Ethier, C., Oby, E. R., Bauman, M. J. & Miller, L. E. Restoration of grasp following paralysis through brain-controlled stimulation of muscles. *Nature* **485**, 368–371 (2012).
- Collinger, J. L. *et al.* High-performance neuroprosthetic control by an individual with tetraplegia. *Lancet* **381**, 557–564 (2013).
- Hochberg, L. R. *et al.* Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* **485**, 372–375 (2012).
- Bouton, C. E. *et al.* Restoring cortical control of functional movement in a human with quadriplegia. *Nature* **533**, 247–250 (2016).
- Mushahwar, V. K., Guevremont, L. & Saigal, R. Could cortical signals control intraspinal stimulators? A theoretical evaluation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **14**, 198–201 (2006).

6. Ho, C. H. *et al.* Functional electrical stimulation and spinal cord injury. *Phys. Med. Rehabilitation Clinics North Am.* **25**, 631–654 (2014).
7. Kapadia, N. *et al.* A randomized trial of functional electrical stimulation for walking in incomplete spinal cord injury: effects on walking competency. *J. Spinal Cord Med.* **37**, 511–524 (2014).
8. Wenger, N. *et al.* Closed-loop neuromodulation of spinal sensorimotor circuits controls refined locomotion after complete spinal cord injury. *Sci. Transl. Med.* **6**, 255ra133 (2014).
9. Wenger, N. *et al.* Spatiotemporal neuromodulation therapies engaging muscle synergies improve motor control after spinal cord injury. *Nat. Med.* **22**, 138–145 (2016).
10. Moraud, E. M. *et al.* Mechanisms underlying the neuromodulation of spinal circuits for correcting gait and balance deficits after spinal cord injury. *Neuron* **89**, 814–828 (2016).
11. Sherrington, C. S. Flexion-reflex of the limb, crossed extension-reflex, and reflex stepping and standing. *J. Physiol.* **40**, 28–121 (1910).
12. Kiehn, O. Decoding the organization of spinal circuits that control locomotion. *Nat. Rev. Neurosci.* **17**, 224–238 (2016).
13. Holinski, B. J., Everaert, D. G., Mushahwar, V. K. & Stein, R. B. Real-time control of walking using recordings from dorsal root ganglia. *J. Neural Eng.* **10**, 056008 (2013).
14. Angeli, C. A., Edgerton, V. R., Gerasimenko, Y. P. & Harkema, S. J. Altering spinal cord excitability enables voluntary movements after chronic complete paralysis in humans. *Brain* **137**, 1394–1409 (2014).
15. Gerasimenko, Y. P. *et al.* Noninvasive reactivation of motor descending control after paralysis. *J. Neurotrauma* **32**, 1968–1980 (2015).
16. Danner, S. M. *et al.* Human spinal locomotor control is based on flexibly organized burst generators. *Brain* **138**, 577–588 (2015).
17. Barthelemy, D., Leblond, H. & Rossignol, S. Characteristics and mechanisms of locomotion induced by intraspinal microstimulation and dorsal root stimulation in spinal cats. *J. Neurophysiol.* **97**, 1986–2000 (2007).
18. Shenoy, K. V. & Carmena, J. M. Combining decoder design and neural adaptation in brain-machine interfaces. *Neuron* **84**, 665–680 (2014).
19. Shانهchi, M. M., Hu, R. C. & Williams, Z. M. A cortical-spinal prosthesis for targeted limb movement in paralysed primate avatars. *Nat. Commun.* **5**, 3237 (2014).
20. Zimmermann, J. B. & Jackson, A. Closed-loop control of spinal cord stimulation to restore hand function after paralysis. *Front. Neurosci.* **8**, 87 (2014).
21. Nishimura, Y., Perlmutter, S. I., Eaton, R. W. & Fetzi, E. E. Spike-timing-dependent plasticity in primate corticospinal connections induced during free behavior. *Neuron* **80**, 1301–1309 (2013).
22. Lemon, R. N. Descending pathways in motor control. *Annu. Rev. Neurosci.* **31**, 195–218 (2008).
23. Friedli, L. *et al.* Pronounced species divergence in corticospinal tract reorganization and functional recovery after lateralized spinal cord injury favors primates. *Sci. Transl. Med.* **7**, 302ra134 (2015).
24. Courtine, G. *et al.* Can experiments in nonhuman primates expedite the translation of treatments for spinal cord injury in humans? *Nat. Med.* **13**, 561–566 (2007).
25. Yin, M. *et al.* Wireless neurosensor for full-spectrum electrophysiology recordings during free behavior. *Neuron* **84**, 1170–1182 (2014).
26. Yakovenko, S., Mushahwar, V., VanderHorst, V., Holstege, G. & Prochazka, A. Spatiotemporal activation of lumbosacral motoneurons in the locomotor step cycle. *J. Neurophysiol.* **87**, 1542–1553 (2002).
27. Rattay, F., Minassian, K. & Dimitrijevic, M. R. Epidural electrical stimulation of posterior structures of the human lumbosacral cord: 2. Quantitative analysis by computer modeling. *Spinal Cord* **38**, 473–489 (2000).
28. Courtine, G. & Bloch, J. Defining ecological strategies in neuroprosthetics. *Neuron* **86**, 29–33 (2015).
29. Shoham, S., Halgren, E., Maynard, E. M. & Normann, R. A. Motor-cortical activity in tetraplegics. *Nature* **413**, 793 (2001).
30. van den Brand, R. *et al.* Restoring voluntary control of locomotion after paralyzing spinal cord injury. *Science* **336**, 1182–1185 (2012).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** G.C. holds the International Foundation for Research in Paraplegia Chair in Spinal Cord Repair. S.M. holds the Bertarelli Foundation Chair in Translational Neuroengineering. We thank X. Rulin and C. Yunlong for providing support, taking care of the monkeys, performing behavioural training, and collecting data; E. Pirondini, N. Pavlova and P. Musienko for help with experiments; J. Courtine, I. Pitteloud, J. Rubattel, L. Dalang and R. Hasler for help with kinematic reconstruction; J. Courtine for the voice-over in video; J. Kreider for help with anatomy; and J. Laurens for discussions and photographs. The illustrations were created by Jemère Ruby. This work was supported by Medtronic, the European Community's Seventh Framework Program (CP-IP 258654, NeuWALK), the International Paraplegic foundation, a Starting Grant from the European Research Council (ERC 261247, Walk Again), the Wyss centre in Geneva, a Marie Curie Fellowship to D.B. (331602, e-WALK), Marie Curie COFUND EPFL fellowships to T.M. and F.W., a Morton Cure Paralysis Fund fellowship to T.M., and the Swiss National Science Foundation including the National Centre of Competence in Research in Robotics, the Sinergia program (CRSII3\_160696), the Sino-Swiss Science and Technology Cooperation (I2L2CZ3\_156331), and the NanoTera.ch programme (SpineRepair).

**Author Contributions** M.C., T.M. and D.B. contributed equally to this work. F.W. and E.M.M. contributed equally to this work. S.M., E.B. and J.B. contributed equally to this work. M.C. developed the spinal cord stimulation protocols and the routines for the identification of flexion and extension hotspots. T.M. developed the brain decoder and the decoder calibration routines. D.B. developed the experimental platform. M.C., T.M., F.W. and E.M.M. performed all the behavioural experiments (with help from D.B., J.G., Y.J. and G.C.). M.C., T.M. and F.W. analysed the data (with input from E.M.M., J.-B.M. and D.X.). M.C., T.M., F.W., E.M.M. and J.G. developed the real-time software application. N.B. and T.D. developed the Neural Research Programmer (with input from M.C., D.B., T.M., F.W. and J.G.). Q.B. and E.R. processed the anatomical data. Y.J. trained all the monkeys. W.K.D.K., Q.L. and E.B. managed the experimental protocols and procedures. P.D. developed and produced the spinal implants (from designs by M.C., D.B., J.B. and G.C.). J.B., D.B., Q.L. and G.C. performed the surgeries. G.C., S.M., E.B., J.B. and P.D. secured funding for the study. G.C. conceived and supervised the study. G.C. wrote the paper with M.C., T.M. and F.W., and all the authors contributed to its editing.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.C. ([gregoire.courtine@epfl.ch](mailto:gregoire.courtine@epfl.ch)).

**Reviewer Information** *Nature* thanks A. Jackson, A. Prochazka, S. Scott and the other anonymous reviewer(s) for their contribution to the peer review of this work.



## METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

**Animal husbandry.** Experiments were approved by the Institutional Animal Care and Use Committee of Bordeaux (CE50, France) under the license number 50120102-A and performed in accordance with the European Union directive of 22 September 2010 (2010/63/EU) on the protection of animals used for scientific purposes in an AAALAC-accredited facility (Chinese Academy of Science, Beijing, China). Nine healthy male rhesus monkeys (*Macaca mulatta*, China; Supplementary Table 1) aged between 4 and 9 years old, and weighing between 4.3 kg and 8.4 kg ( $6.5 \pm 0.5$  kg) were housed individually in cages designed according to European guidelines ( $2 \text{ m} \times 1.6 \text{ m} \times 1.26 \text{ m}$ ). Environmental enrichment included toys and soothing music. All the monkeys are included in the manuscript. Only two monkeys received a spinal cord injury.

**Surgical procedures.** All the surgical procedures were performed under full anaesthesia induced with atropine ( $0.04 \text{ mg kg}^{-1}$ ) and ketamine ( $10 \text{ mg kg}^{-1}$ , intramuscular injection) and maintained under 1%–3% isoflurane after intubation. A certified functional neurosurgeon (J.B.) supervised all the surgical procedures. Surgical implantations were performed during a single operation lasting approximately 8 h. We implanted a 96-channel microelectrode array (Blackrock Microsystems, of pitch 1.5 mm) into the leg area of the left primary motor cortex<sup>1</sup> (F4, Supplementary Table 1). The monkeys also received a wireless system<sup>2</sup> (T33F-4, Konigsberg Instruments, USA) to record electromyographic signals from the following leg muscles: gluteus medius (GLU), iliopsoas (IPS), rectus femoris (RF), semitendinosus (ST), gastrocnemius medialis (GM), tibialis anterior (TA), extensor digitorum longus (EDL), and flexor hallucis longus (FHL). A custom-made spinal implant was inserted into the epidural space of the lumbar spinal cord according to previously described methods<sup>9</sup>. The implant was inserted at the L4–L5 vertebrae and pulled until it reached the T13–L1 vertebrae. Electrophysiological testing was performed intra-operatively to adjust the position of the electrodes. Specifically, we verified that a single pulse of stimulation delivered through the most rostral and most caudal electrodes induced motor responses in the IPS and GM muscles, respectively. The connector of the implant, enclosed into a titanium orthosis, was secured to the vertebral bone using titanium screws (Vis MatrixMIDFACE, of diameter 1.5 mm and length 8 mm, Synthes). The wires were routed subcutaneously to an implantable pulse generator inserted between intercostal muscles (see Supplementary Information).

Monkeys Q2 and Q3 received a spinal cord injury. A partial laminectomy was made at the level of the T7/T8 thoracic vertebrae. A micro-blade was used to cut approximately two-thirds of the dorsoventral extent of the spinal cord. The lesion was completed using micro-scissors under microscopic observation. Animals retained bowel, bladder, and autonomic function after the injury. The veterinary team continuously monitored the monkeys during the first hours after surgery, and numerous times daily during the seven subsequent days. A few hours after completion of surgical interventions, the animals were able to move around and feed themselves unaided. Clinical rating and monitoring scales were used to assess post-operative pain. Ketophen ( $2 \text{ mg kg}^{-1}$ ; subcutaneous) and Metacam ( $0.2 \text{ mg kg}^{-1}$ ; subcutaneous) were administered once daily. Lidocaine cream was also applied to surgical wounds twice per day. The antibiotics ceftriaxone sodium ( $100 \text{ mg kg}^{-1}$ ; intramuscular) was given immediately following surgery, and then once daily for 7 days.

**Experimental recordings.** Monkeys were trained to walk on a treadmill and overground along a corridor ( $300 \text{ cm} \times 35 \text{ cm} \times 70 \text{ cm}$ ). Plexiglas enclosures were used to maintain the monkeys within the field of view of the cameras. Food pellets and fruits rewarded appropriate behaviours. Additional food to complete daily dietary requirements was provided after training.

**Single-pulse stimulation in sedated monkeys.** Monkeys were lightly sedated with ketamine ( $3.5 \text{ mg kg}^{-1}$ ), and suspended in the air using a jacket that did not impede leg movements. Single pulses of cathodic monopolar, charge-balanced stimulation ( $0.3 \text{ ms}$ ,  $1 \text{ Hz}$ ) were delivered through the electrodes to elicit compound potentials in leg muscles. We selected the active sites whose corresponding spatial maps of motoneuron activation showed the highest correlation with the hotspots.

**Brain-controlled stimulation during locomotion in intact monkeys.** Brain-controlled stimulation protocols were tested during locomotion on a treadmill at a comfortable speed (Q1,  $2.0 \text{ km h}^{-1}$ ; Q2,  $1.6 \text{ km h}^{-1}$ ). Recording sessions were organized as follows: first, we recorded two to five blocks each of duration 1–2 min during stepping without stimulation. These baseline recordings were used to calibrate the decoders for real-time detection of foot-off and foot-strike gait events. Second, monkeys were recorded during brain-controlled stimulation protocols involving (1) solely the electrode targeting the extensor hotspot, (2) solely the electrode targeting the flexor hotspot, and (3) both electrodes. We tested the

effects of stimulation frequency and amplitude over functional ranges ( $30$ – $80 \text{ Hz}$ ;  $1.5$ – $3.9 \text{ V}$ ). See Supplementary Table 2.

**Brain-controlled stimulation during locomotion in injured monkeys.** Monkeys Q2 and Q3 were recorded after injury as soon as they were able to sustain independent locomotion on the treadmill, which corresponded to 6 days and 16 days post-injury, respectively. Q3 recovered more slowly than Q2, probably owing to receiving more extensive ventral and lateral spinal cord damage (Extended Data Fig. 9). Therefore, monkey Q3 could only be recorded when appropriate behavioural and physical conditions were reached, which occurred two weeks post-injury. Owing to restrictions on the total duration of the experiments (2 weeks), only one entire session could be conducted with this monkey. Following this experiment, the monkey rapidly recovered, which prevented evaluation of the efficacy of the brain–spine interface. The monkeys were recorded on the treadmill at their most comfortable speed ( $1.2$ – $1.4 \text{ km h}^{-1}$  for monkey Q2 and  $1.0 \text{ km h}^{-1}$  for monkey Q3). Recording sessions were organized as follows. First, we recorded two to six blocks each of duration 1–2 min without stimulation. These recordings were used to calibrate the decoders. Second, the decoders were used to test brain-controlled stimulation of both the extension and flexion hotspots over a range of stimulation frequencies. The effects of continuous stimulation using the same stimulation features as during brain-controlled stimulation were also tested. Within the functional range of stimulation parameters, brain-controlled stimulation did not trigger undesired movements or spasms that impaired locomotor movements. See Supplementary Table 2.

**Data acquisition.** Procedures to record kinematics and muscle activity have been detailed previously<sup>25,31</sup>. Whole-body kinematics was measured using the high-speed motion capture system SIMI (Simi Reality Motion Systems, Germany), combining 4 or 6 video cameras ( $100 \text{ Hz}$ ). Reflective white paint was directly applied on the shaved skin of the monkey overlying the following body landmarks of the right side: iliac crest, greater trochanter (hip), lateral condyle (knee), lateral malleolus (ankle), 5th metatarsophalangeal (mtp), and the outside tip of the fifth digit (toe). The Simi motion tracking software was used to obtain the three-dimensional spatial coordinates of the markers. Joint angles were computed accordingly. Electromyographic signals were recorded simultaneously ( $2 \text{ kHz}$ , Konigsberg, USA) and synchronized through the Blackrock Cerebrus system (Blackrock Microsystems, USA), which also recorded neural signals. For this, the Cereplex wireless system<sup>25</sup> was mounted on the head of the monkeys. Six antennae and a receiver were used to transmit<sup>25</sup> broadband neural signals ( $0.1 \text{ Hz}$  to  $7.8 \text{ kHz}$  band, sampled at  $22 \text{ kHz}$ ). The signals were band-pass filtered ( $500 \text{ Hz}$  to  $7.5 \text{ kHz}$ ) and spiking events were extracted through threshold crossings<sup>2,3,32–34</sup>. Specifically, a spiking event was defined on each channel (96 in total) if the signal exceeded  $3.0$ – $3.5$  times its root-mean-square value calculated over a period of  $5 \text{ s}$ . This procedure resulted in a binary signal from 96 multiunits, each originating from one of the 96 electrodes of the array. Signals from all 96 multiunits were integrated in the decoder.

**Decoding of motor states from neural signals.** Our aim was to deliver stimulation over the extensor and flexor hotspots around the times at which these hotspots are active during natural locomotion. To this end, we decoded gait-related motor states from neural activity and used those detections to trigger the stimulation protocols at the appropriate times. The control computer was connected to the local network and continuously received Used Datagram Protocol (UDP) packets containing neural recordings. We designed a custom in-house software application running on the control computer (Visual Studio C++ 2010), which analysed the neural signals in real time. Every  $20 \text{ ms}$ , the application made a decision: whether or not to trigger one of the spinal cord stimulation protocols. The decision was made based on probabilities of observing a foot-off or a foot-strike motor state given the history of neural data ( $300 \text{ ms}$  pre-lesion and  $400 \text{ ms}$  post-lesion), as calculated by our decoders.

Natural activations of the extension and flexion hotspots were time-locked to foot-off and foot-strike gait events, respectively). In turn, we defined the foot-off and foot-strike motor states as the neural activity preceding foot-off and foot-strike gait events by  $\Delta t_{\text{FO}}$  and  $\Delta t_{\text{FS}}$  temporal offsets, respectively. The offsets were derived to maximize the overlap between the stimulation over the hotspots and the natural activation of those hotspots. In effect, the offsets integrated the latencies between the gait events and the hotspot activations, as well as latencies related to wireless communication between our devices, into the design of our decoders.

**Extraction of motor states used for decoder calibration.** We calibrated the decoders on data from two to seven no-stimulation blocks recorded at the beginning of each session. Gait events were identified from electromyographic recordings (Q1) or from video recordings (Q2 and Q3). Identification of foot-off and foot-strike gait events from electromyographic recordings was performed using signals from the iliopsoas muscle, which was active around the time of swing onset and remained active throughout most of the swing phase of gait. The foot-off and foot-strike

gait events were estimated by thresholding the envelope of the rectified electromyographic signal. Identification of foot-off and foot-strike gait events from video recordings was performed visually. After injury and while the monkeys exhibited minimal movements only of the limb ipsilateral to lesion, foot-off and foot-strike gait events were defined according to residual hip or knee oscillations, which correlated with attempts to execute steps.

**Calibration to account for stimulation-induced changes in neural signals.** Analysis of the decoding temporal precision in Q1 revealed that decoded foot-off and foot-strike motor states during brain-controlled stimulation differed from the times of the motor states estimated from the foot-off and foot-strike gait events (median difference for foot-off was 68 ms and for foot-strike was -90 ms). We did not observe such a difference when detecting motor states in the absence of stimulation (median difference for foot-off was 11 ms and for foot-strike was 3 ms). A range of factors could have decreased decoding performance, including changes in somatosensory feedback influenced by the stimulation, attempts by a monkey to adapt its gait, changes in stability, and so on. To improve the temporal accuracy of our decoder, we introduced a decoder recalibration process. The initial decoder, trained on data without stimulation, was used to trigger stimulation through the extension hotspot or flexion hotspot independently for 2 to 3 blocks each. The data collected during these blocks was then combined with the blocks of no stimulation to calibrate a new, second decoder. This decoder successfully compensated for stimulation-induced changes in motor cortex activity (Fig. 3, and Extended Data Fig. 5).

**Duration of hotspot stimulation protocols.** We sought to stimulate flexion and extension hotspots throughout the duration of their natural activation during locomotion. We determined the duration of the flexion and extension hotspot stimulation protocols by setting this duration to 300 ms. We then recorded a few steps during brain-controlled stimulation, and adjusted the duration of the stimulation protocols for each monkey when necessary in order to obtain a clear modulation of leg kinematics. This procedure was performed only once for all pre-injury sessions and was repeated for each post-injury session.

**Blinding.** Data analyses, except for identification of the steps and the marking of foot-off and foot-strike gait events from video recordings, were performed by automatic computer routines. When analyses required involvement of investigators, these were blind to the experimental conditions.

**Spatiotemporal map of motoneuron activation.** To visualize spatiotemporal maps of motoneuron activation, electromyographic signals were mapped onto the rostrocaudal distribution of the motoneurons reconstructed from histological analyses. This approach provides an interpretation of the motoneuron activation at a segmental level rather than at the individual muscle level.

**Identification of extensor and flexor hotspots activation.** Flexion and extension hotspots were identified from the mean spatiotemporal map of motoneuron activation for each monkey independently ( $n = 3$  for Q1, P2 and P3). Single maps computed between two consecutive foot-strike events were time-interpolated to a 1,000-point map and averaged to obtain the mean spatiotemporal map of motoneuron activation. Flexion and extension hotspots were then identified by time-averaging the mean map around the foot-off event ( $-10\% + 20\%$  of the gait cycle) for the flexion hotspot and around the foot strike event ( $-10\% + 30\%$  of the gait cycle) for the extension hotspot.

**Analysis of muscle recruitment curves.** The compound potentials recorded in leg muscles were rectified and integrated for each muscle and stimulation amplitude, and represented in colour-coded spatial maps of motoneuron activation. Instead of measuring specific flexor and extensor muscle selectivity we selected the electrodes that elicited spatial maps similar to those extracted during activation of the flexion and extension hotspots, regardless of muscle specificity. The correlation between the resulting map and the maps recorded during locomotion was calculated for each monkey to identify the voltage range over which the correlation was maximal. The derived voltage range was then used during behavioural experiments (Extended Data Fig. 3c).

**Decoding performance quantification.** We quantified the performance of our asynchronous decoders using confusion matrices and normalized mutual information, as described in ref. 35.

**Steps classification for kinematic analysis.** To evaluate the efficacy of the brain-spine interface and assess the importance of the timing of stimulation in correcting gait deficits, we conducted a post-hoc classification of the steps based on the temporal accuracy of the decoder to reproduce the desired hotspot activation timings. We defined optimal and sub-optimal steps according to the initiation

of flexion and extension hotspot stimulation. All the gait cycles that contained only one correct extension activation (stimulation occurring at foot-strike  $\pm 125$  ms) and only one correct flexion activation (stimulation occurring between foot-off -200 ms and foot-off +50 ms) were defined as optimal steps (Extended Data Fig. 10).

**Stepping quantity.** After the spinal cord lesion, the monkeys typically walked using their three intact limbs while the leg ipsilateral to the lesion was either dragging along the walking surface or maintained in a flexed posture. Occasionally, monkeys hopped to move both legs forward and avoid bumping against the back of the treadmill enclosure, owing to their inability to move at the selected treadmill belt speed. We counted the numbers of these 'hop' and 'bump' steps as well as the numbers of normal steps. Experimenters were blinded to stimulation conditions during this analysis. To quantify the functional improvement mediated by the brain-spine interface, we calculated the proportion of normal steps over all recorded blocks on a given day. To quantify the ability of the monkeys to sustain locomotion, we extracted all the events marked as steps, and measured the relative number of steps that were not performed while bumping into the back of the treadmill enclosure.

**Stepping quality.** A total of 26 parameters quantifying kinematics (Supplementary Table 3) were computed for each step according to methods described previously in refs 8, 9 and 31. We used principal component analysis to visualize the changes in gait over time and for different conditions (Fig. 4, and Extended Data Figs 7 and 8). To quantify locomotor performance, we calculated the mean Euclidean distance between steps corresponding to a given experimental condition and the mean of steps recorded before the lesion in the same monkey in the entire 26-dimensional space of kinematic parameters.

**Tissue processing.** Monkeys were deeply anesthetized and perfused transcardially with a 4% solution of paraformaldehyde. The spinal cord dura was removed and the spinal cord was removed by cutting 40- $\mu$ m-thick sections using a cryostat microtome, before storage at 4°C in 0.1 M phosphate-buffered saline azide (0.03%).

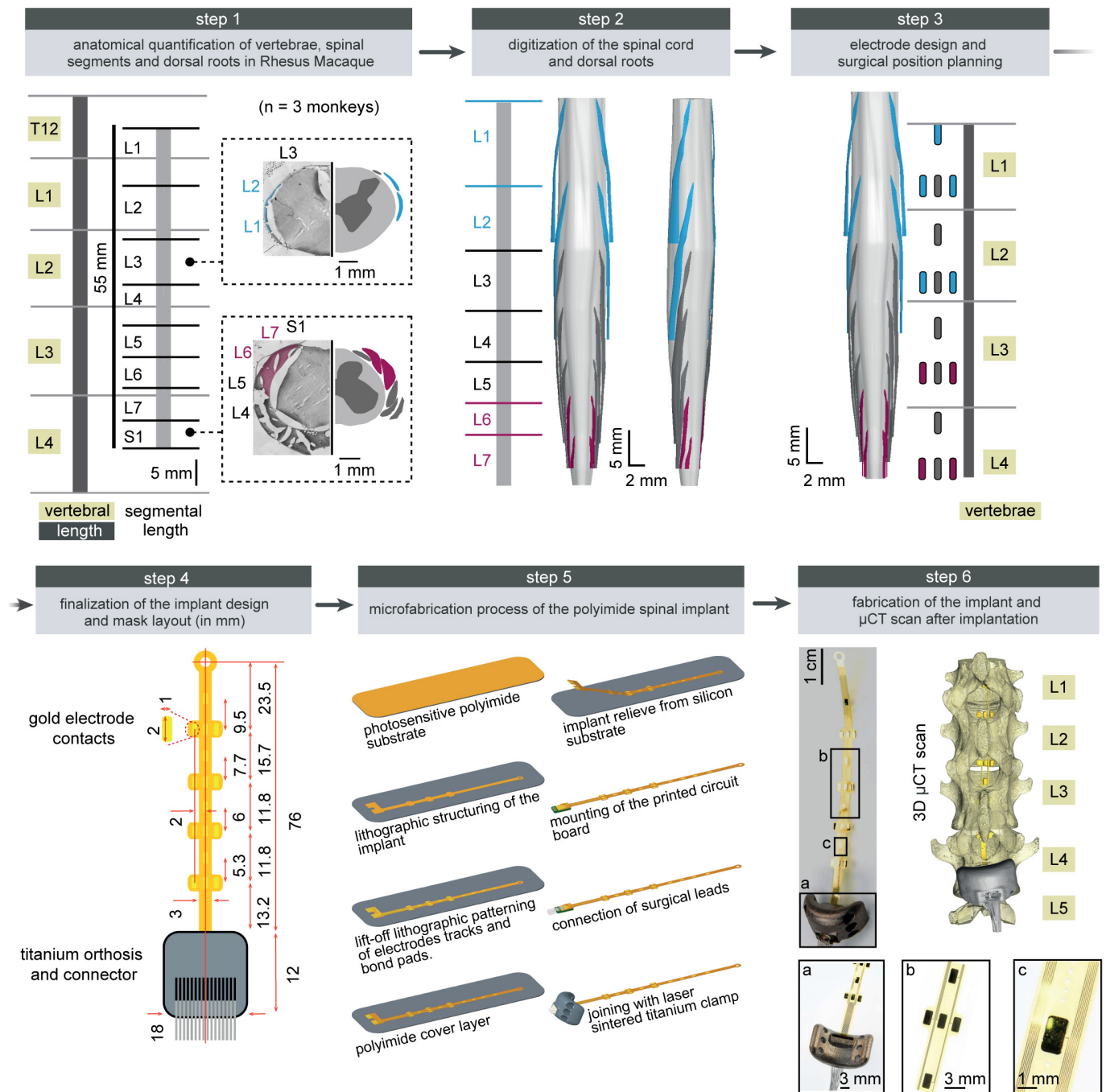
**Anterograde tracing of motor cortex projections.** Monkeys Q2 and Q3 underwent anterograde tracing of corticospinal projections from the leg and trunk area of the left motor cortex using anatomical tracers. All animals were anesthetized as described above. Biotinylated dextran amine (BDA; 10% solution in water; 10,000 Da; Molecular Probes, TSA PLUS Biotin KIT PerkinElmer, catalogue number NEL749A001KT) was injected at 300 nl per site into 40 sites spanning the leg and trunk regions of the left motor cortex.

**Quantification of the spinal cord lesion.** Camera lucida reconstructions of the lesion (NeuroLucida 11.0, MBF Biosciences, USA) were performed using evenly spaced horizontal sections (1:4) throughout the whole dorsoventral axis on sections labelled for astrocytic (GFAP; 1:1000, Dako, USA, catalogue number Z0334), NeuN (anti-NeuN; 1:300, Millipore, catalogue number MAB377) and BDA reactivity. Immunoreactions were visualized with secondary antibodies labelled with AlexaFluor 488 (1:400, Invitrogen, catalogue number A-11034) and 647 (1:300, Invitrogen, catalogue number A-21235).

**Statistical procedures.** All the computed parameters were quantified and compared within each monkey. All data are reported as mean values  $\pm$  standard error of the mean (s.e.m.). Significance was analysed using the non-parametric Wilcoxon rank-sum test, bootstrapping or a Monte Carlo approach.

**Data availability.** Data that support the findings and software routines developed for the data analysis are available from the corresponding author upon reasonable request.

- Courtine, G. *et al.* Kinematic and EMG determinants in quadrupedal locomotion of a non-human primate (*Rhesus*). *J. Neurophysiol.* **93**, 3127–3145 (2005).
- Fraser, G. W., Chase, S. M., Whitford, A. & Schwartz, A. B. Control of a brain-computer interface without spike sorting. *J. Neural Eng.* **6**, 055004 (2009).
- Gilja, V. *et al.* Clinical translation of a high-performance neural prosthesis. *Nat. Med.* **21**, 1142–1145 (2015).
- Aflalo, T. *et al.* Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. *Science* **348**, 906–910 (2015).
- Milekovic, T., Ball, T., Schulze-Bonhage, A., Aertsen, A. & Mehring, C. Detection of error related neuronal responses recorded by electrocorticography in humans during continuous movements. *PLoS One* **8**, e55235 (2013).
- Marieb, E. N. *Human Anatomy and Physiology* 6th edn (Pearson Education, 2003).
- Kuypers, H. G. J. M. in *Comprehensive Physiology* Ch. 3 (John Wiley & Sons, 2011).

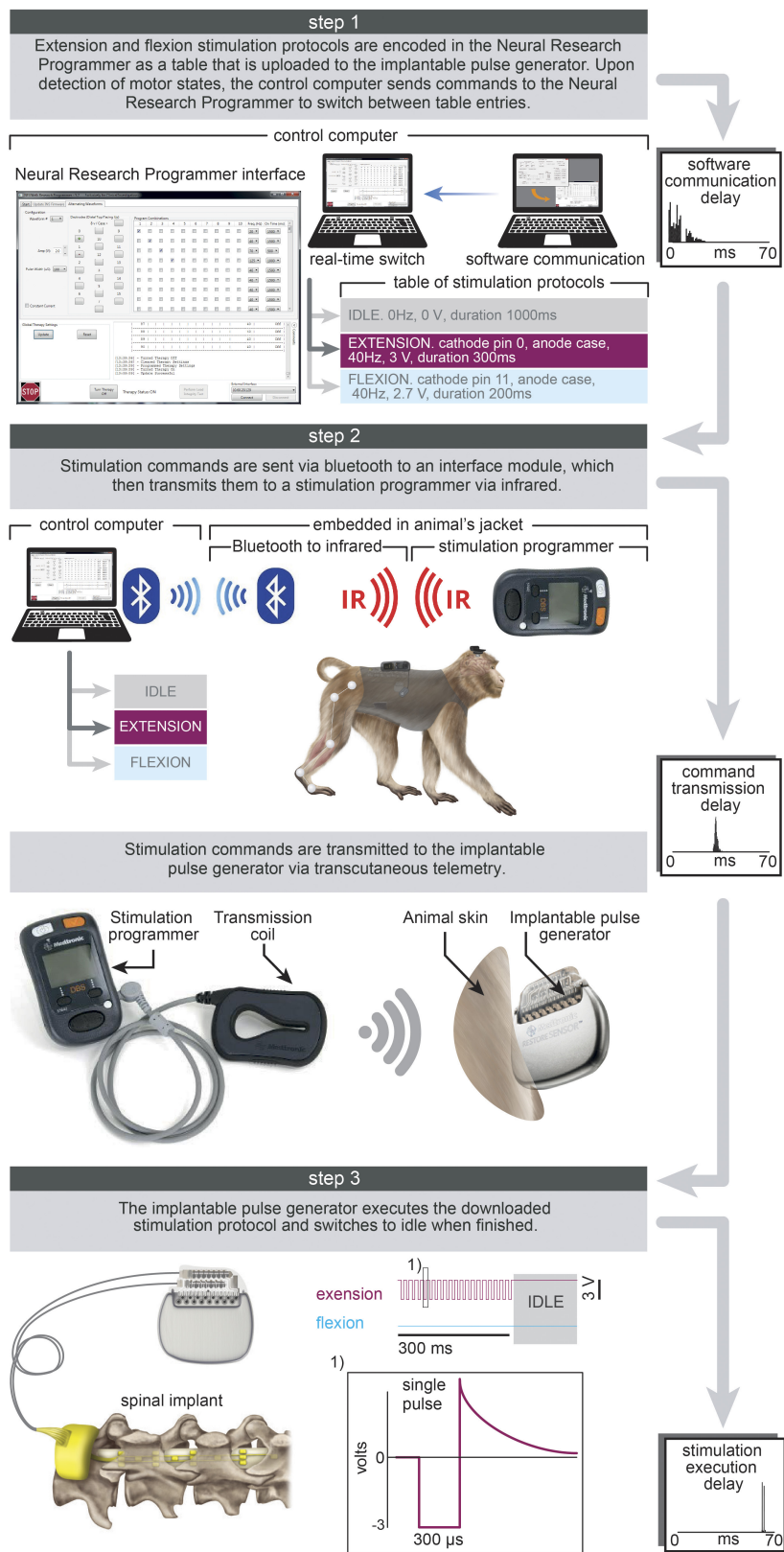


Extended Data Figure 1 | See next page for caption.



**Extended Data Figure 1 | Design and fabrication of the spatially selective spinal implant.** Step 1: Quantification of the length of vertebrae T12 to L4, and of spinal segments L1 to S1. The photographs show coronal sections of the left halves of the L3 and S1 spinal segments, including the dura mater, dorsal roots and spinal tissue. The complementary views show camera lucida reconstructions of the dorsal roots, white and grey matter. The dorsal roots projecting to the L1/L2 and L6/L7 spinal segments are colour-coded to visualize their respective location. Step 2: Three-dimensional reconstruction of the entire lumbosacral spinal cord and dorsal roots. The dorsal roots innervating spinal segments containing the targeted extension (L6, L7) and flexion (L1, L2) hotspots are colour-coded to help visualize their respective spatial trajectory. The three-dimensional reconstructions are displayed from a dorsal view, and from a view that is rotated 45° leftward around the rostrocaudal axis to visualize the trajectory of the dorsal roots along spinal segments. Step 3: Design of the spatially selective spinal implants, including the location of the electrodes with respect to the dorsal roots, and positioning of the implants with respect to the vertebrae. The cyan and magenta shapes highlight the electrodes targeting the flexor and extensor hotspots, respectively. Step 4: Mask layout of the spatially selective epidural spinal implant. Step 5: Fabrication of the spinal implants. Processing starts with deposition of a 40- $\mu\text{m}$ -thick polyimide film by spin coating on the silicon substrate. The polyimide-based bottom layer constituting the footprint of the implant is realized

through ultraviolet lithography. A 200-nm-thick gold layer is deposited using vacuum evaporation and lithographically structured to create a conductive seed pattern. An additional gold layer is electroplated to a height of approximately 6  $\mu\text{m}$  to create the electrodes and interconnects. A 20- $\mu\text{m}$ -thick cover layer of photosensitive polyimide is then applied by spin coating to uniformly cover the bottom and electrode structure. A final ultraviolet lithography is applied to structure the top layer of the implant, and to create openings over the electrodes and contact pads. The implants are gently released from the carrier wafer. To create a connector, the contact pads located at the extremity of the implant are glued to a small support plate made of medical-grade poly-methyl methacrylate (PMMA) sheet material. Precision milled alignment structures allow fine adjustment of the support plate to the spinal implant, and to the extremity of the leads. The support plate contains a seat in which a tiny printed circuit board with conductive stripes is fixed. Stainless steel leads with silicone rubber insulation are directly soldered to the contact pads of the implant via the contact stripes located on the printed circuit board. The resulting implants and connective leads were highly reliable mechanically and electrically. A fixation structure was realized by using three-dimensional laser sintering of medical grade titanium. Step 6: Photographs, including zoomed insets, showing a fabricated spinal implant and a micro-computed tomography scan performed after implantation. Photographs, copyright J. Laurens (2016).

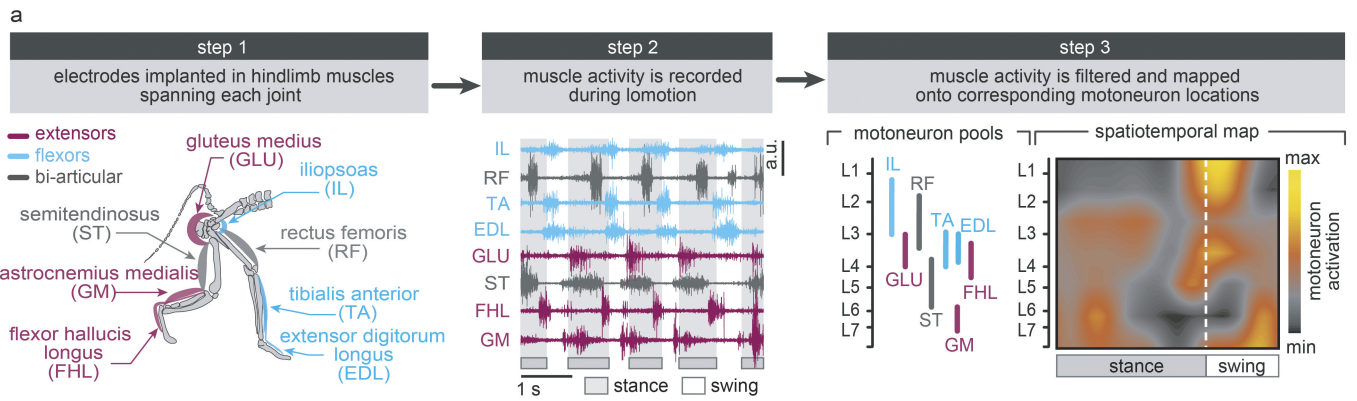


Extended Data Figure 2 | See next page for caption.

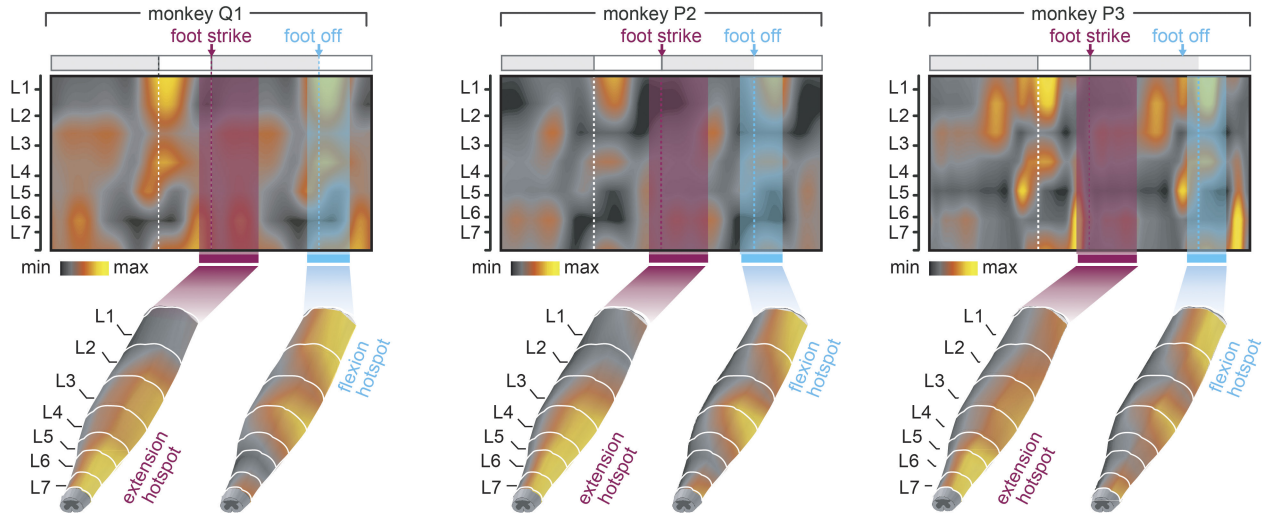
**Extended Data Figure 2 | Protocols and technology of the spinal cord stimulation system.** Step 1: A Neural Research Programmer interface (screen snapshot) encodes stimulation protocols that are pre-programmed into a table uploaded to the implantable pulse generator. Each row of this table corresponds to a specific electrode configuration (cathodes and anodes) and stimulation features (amplitude, frequency, pulse width and duration of stimulation). During experiments, the control computer selects the rows to be executed. The plot reports the distribution of temporal delays introduced by the communication between the decoder and the Neural Research Programmer ( $n = 5,000$ ). Step 2: Stimulation commands are transmitted to the implantable pulse generator. Commands are first broadcast via Bluetooth to a module that converts them into infrared signals transferred to the stimulation programmer device. The Bluetooth-to-infrared module and the stimulation programmer were

embedded into a jacket worn by the monkeys during the experiments. The stimulation programmer transmitted the stimulation commands into the implantable pulse generator via induction telemetry. The antenna was placed under the jacket, in contact with the skin and aligned to the implantable pulse generator. The plot reports the distribution of delays needed to transmit the stimulation commands from the Neural Research Programmer to the implantable pulse generator. Step 3: The implantable pulse generator executed the selected stimulation protocols. After execution of the stimulation command, the implantable pulse generator switched to idle mode. The shape of a single charge balanced cathodic pulse is shown in the inset. The plot reports the distribution of time delays required to execute a single stimulation command by the implantable pulse generator. Photographs, copyright Medtronic (2016); drawing, copyright Jemère Ruby (2016).

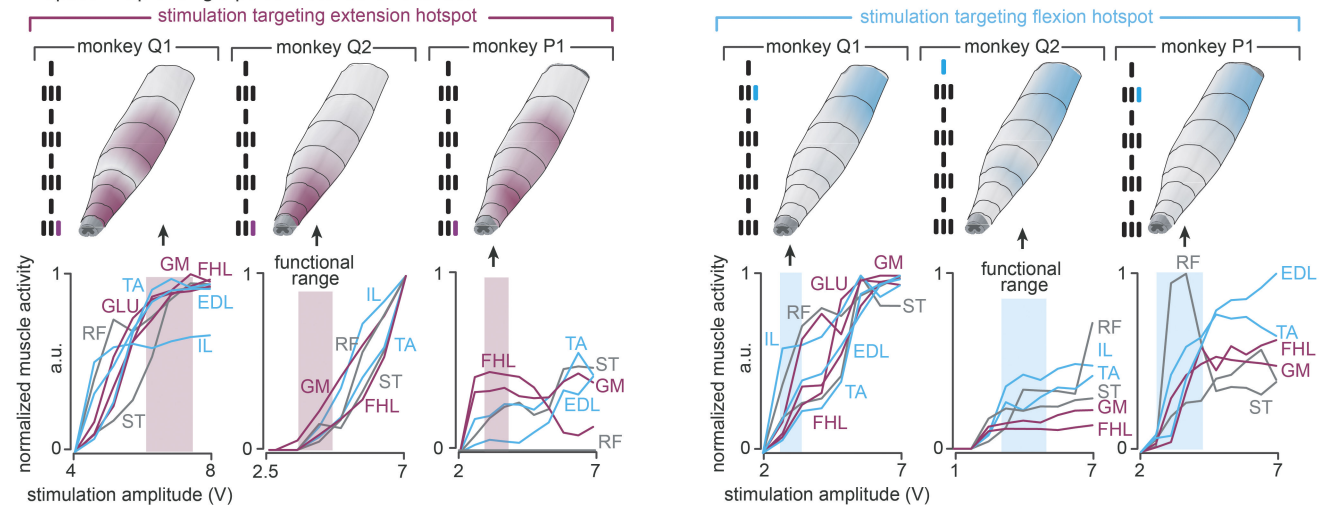




**b** spatiotemporal map of motoneuron activation during locomotion on a treadmill in rhesus macaques



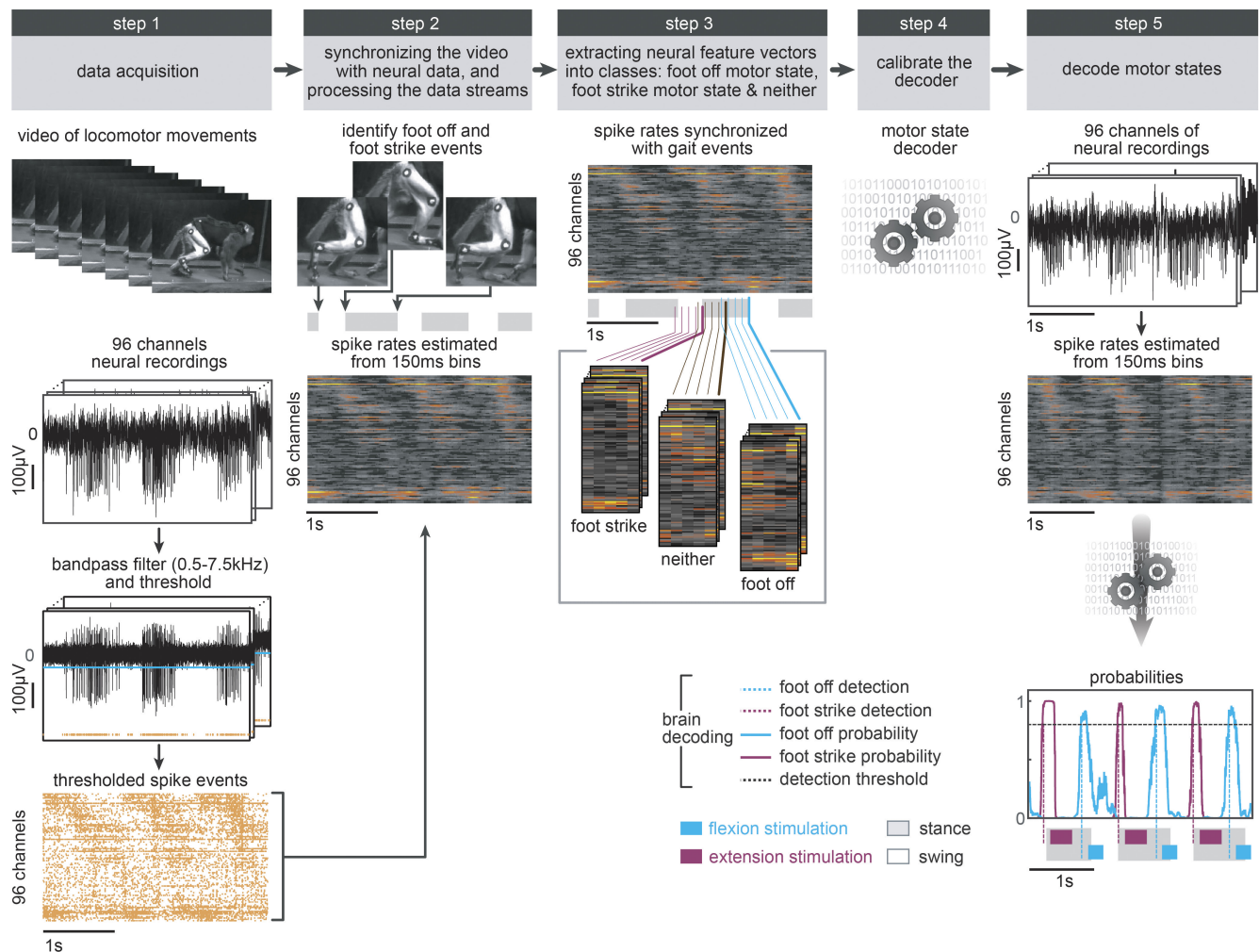
**c** spatial map of single pulse motoneuron activation



Extended Data Figure 3 | See next page for caption.

**Extended Data Figure 3 | Anatomical, computational and functional experiments allowed the identification of stimulation protocols to access flexion and extension hotspots.** **a**, Computational procedure to estimate spatiotemporal maps of motoneuron activation during locomotion. Step 1: Four pairs of antagonist muscles spanning each joint of the leg are implanted with bipolar electrodes to record electromyographic signals during locomotion. Step 2: Muscle activity recorded during locomotion on a treadmill is band-pass-filtered using a Butterworth 3rd-order filter (30–800 Hz, monkey P3). Step 3: The signals are rectified, filtered with a low pass at 10 Hz, normalized to the maximum activity recorded across all the gait cycles, and then projected onto the location of the corresponding motoneuron columns in the spinal cord for each of the recorded muscles. The estimated motoneuron activation is represented as a colour-coded spatiotemporal map of motoneuron activation. **b**, Spatiotemporal maps of motoneuron activation recorded in three intact monkeys (Q1, P2 and P3). The maps were obtained by averaging electromyographic signals recorded during continuous locomotion on a treadmill ( $n = 73, 25$  and  $24$  steps for monkeys Q1, P2 and P3, respectively). The maps underlying the activation of extension and flexion hotspots were extracted by averaging the estimated motoneuron activation around the foot-strike and foot-off events, respectively. For this, a window was defined from  $-10\%$  to  $+30\%$  of the gait cycle duration for the foot-strike event, and from  $-10\%$  to  $+20\%$  of the gait cycle duration

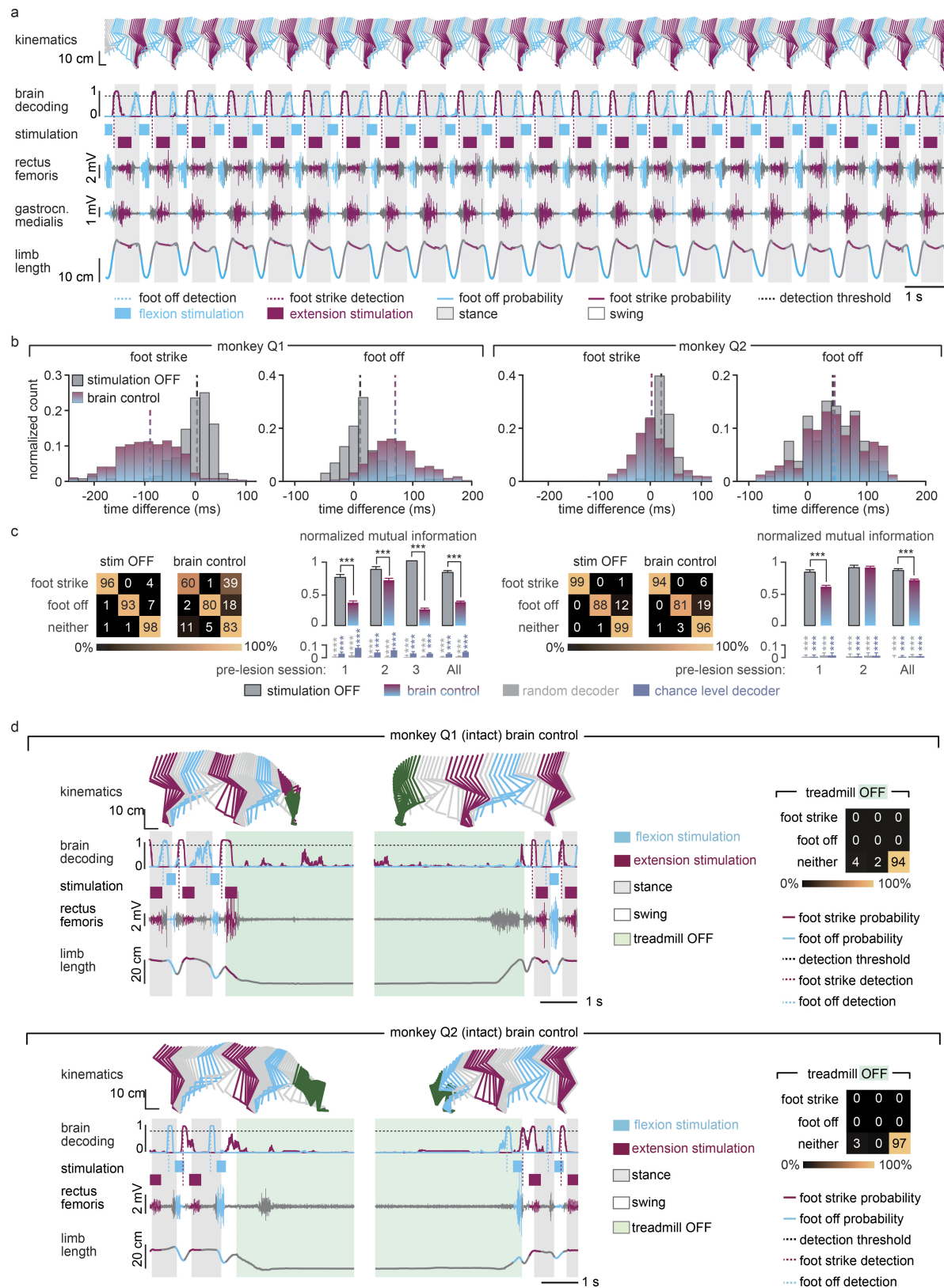
for the foot-off event. The maps were reproducible across monkeys. The correlation between monkeys Q1, P2 and P3 for the flexion hotspot was  $0.94, 0.90$  and  $0.90$  for Q1–P2, Q1–P3 and P2–P3, respectively. The correlation between monkeys Q1, P2 and P3 for the extension hotspot was  $0.88, 0.90$  and  $0.60$  for Q1–P2, Q1–P3 and P2–P3, respectively. The resulting maps were projected onto the reconstructed spinal segments (Extended Data Fig. 1). **c**, Recruitment curves showing the relationships between motor-evoked potentials elicited by single pulses of epidural electrical stimulation in each of the recorded hindlimb muscles and the stimulation amplitude for three intact monkeys (Q1, Q2 and P1). Stimulation was delivered through the electrodes targeting the extension and flexion hotspots. The compound responses elicited in leg muscles were rectified and integrated to calculate the amplitude of the responses, and then projected on the reconstructed spinal segments. The spatial maps of motoneuron activation corresponding to the optimal range of stimulation amplitudes to stimulate the hotspots are displayed for each monkey, including the location of the electrodes with respect to spinal segments. To compute the optimal range of stimulation amplitudes for each hotspot, we extracted the stimulation amplitudes for which the spatial map of motoneuron activation displayed the highest values of correlation with the spatial maps of the targeted hotspots. The cyan and magenta shadings highlight the functional range of stimulation amplitude for each hotspot and monkey. (a.u., arbitrary units.)



**Extended Data Figure 4 | Procedure to calibrate the decoders for real-time detection of motor states.** Step 1: The locomotor movements of the right leg were recorded using the video camera system. In parallel, neural signals were recorded from the microelectrode array implanted into the leg area of the primary motor cortex. The signals were band-pass filtered (0.5–7.5 kHz). A threshold was at 3–3.5 times the standard deviation in order to obtain spike events. The two data streams were saved onto computers. Step 2: Visual inspection of the video frames allowed the identification of foot-off and foot-strike gait events. We estimated the spike rates from overlapping 150-ms bins that were updated every 20 ms. Gait events were then synchronized with the spike rate estimates using a trigger saved with the neural data that marked the onset of video recordings. Step 3: We extracted feature vectors that originated at foot-off and foot-strike events and assigned them to ‘foot-off’ and ‘foot-strike’

motor state classes, respectively. All other feature vectors were assigned to the ‘neither’ class. Step 4: Motor state classes of feature vectors were used to calibrate a regularized linear discriminant analysis decoder. Step 5: The decoder was uploaded into our real-time analysis application running on the control computer. Neural data was collected in real time, processed into spike-rate estimates, and passed through the decoder that calculated the probabilities of foot-off and foot-strike motor states. When one of the motor state probabilities crossed a threshold of 0.8, a command to trigger the flexion or extension hotspot stimulation protocols was sent to the Neural Research Programmer, which relayed this instruction to the implanted pulse generator. Owing to the wireless communication, the command was executed 178 ms (Q1) or 105 ms (Q2 and Q3) after the detection of the motor states.

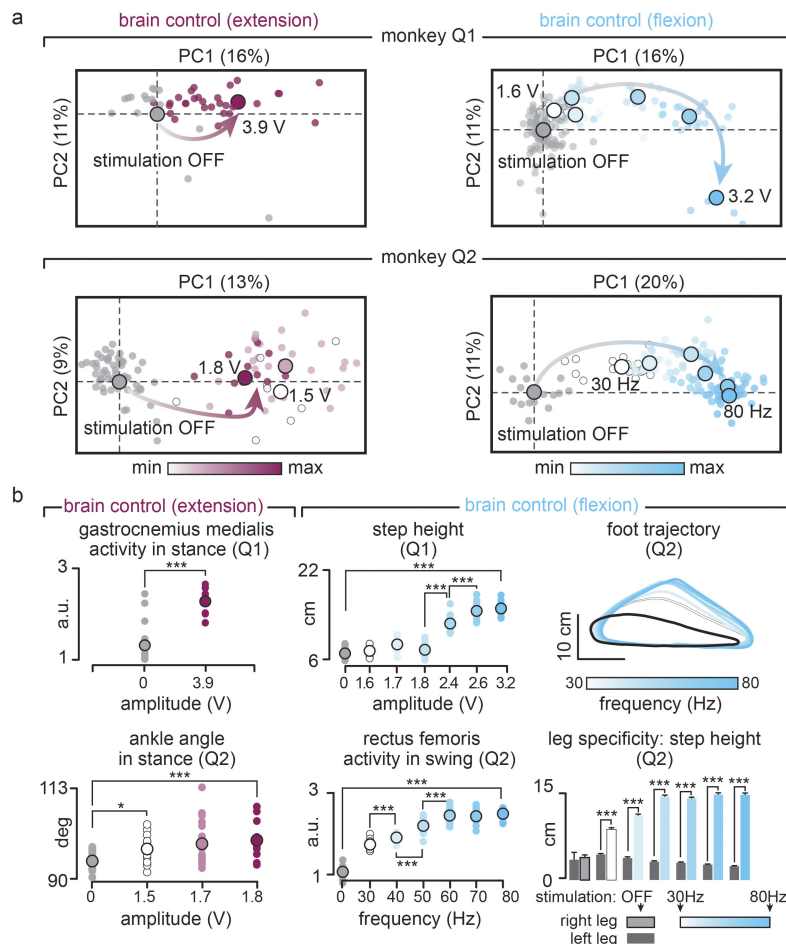




Extended Data Figure 5 | See next page for caption.

**Extended Data Figure 5 | The real-time decoder accurately detected the motor states and triggered stimulation protocols during locomotion and when initiating and terminating it in intact monkeys.** **a**, Example of a continuous sequence of locomotion (20 s) during brain-controlled flexion and extension stimulation in an intact monkey (Q2, pre-lesion session 2; see Supplementary Table 2). Conventions are the same as in Fig. 3. The real-time decoder correctly detected the succession of 22 foot-off and foot-strike motor states occurring in this sequence (vertical dotted lines), and appropriately triggered the relevant stimulation protocols throughout the locomotor sequence. **b**, Histograms showing the distribution of the temporal differences between the actual occurrence of foot-off and foot-strike events and the decoded occurrence of these motor states for all the recording sessions of the tested intact monkeys (Q1 and Q2). The dotted lines indicate the median for each distribution. For monkey Q1, the decoders were calibrated using recordings without stimulation only. For monkey Q2, we improved the accuracy of the decoders by performing the calibration twice. First, the decoders were recalibrated using recordings without stimulation. These decoders were used to collect recordings during brain-controlled flexion or brain-controlled extension independently. New decoders were then calibrated using all the recording blocks together. **c**, Confusion matrices reporting

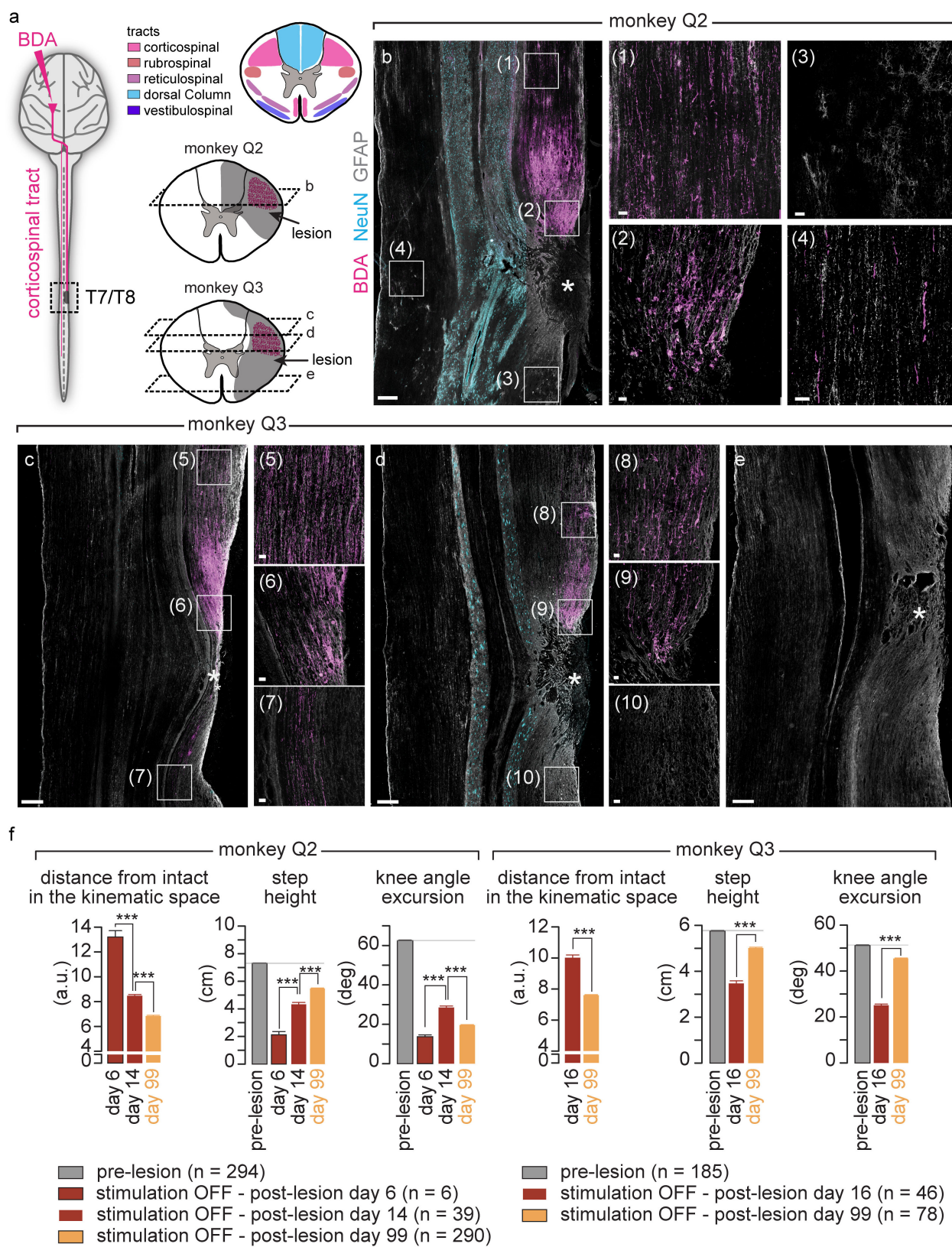
the accuracy of the real-time decoders without stimulation and during brain-controlled stimulation, given a tolerance window of  $\pm 125$  ms. The bar plots report the normalized mutual information calculated for the real-time decoders compared to random decoders and chance level decoders. Random decoders were the same decoders calibrated using shuffled gait events. The chance level decoders randomly assigned motor states with chance level probabilities, which were estimated from the data used to calibrate the online decoders, and at the same rate as the online decoders.  $***P < 0.001$ , bootstrap. Error bars, s.e.m. **d**, Two examples of recordings collected while monkey Q1 and Q2 initiated or terminated sequences of continuous locomotion on a treadmill. The green shaded area indicates the period during which the treadmill was turned off. The brain–spine interface was kept on throughout the recordings. Conventions are the same as in Fig. 3. The probability of foot-strike and foot-off motor states remained low during the periods of rest, and recommenced to modulate with the occurrence of motor states when the monkey resumed continuous locomotion. The decoder confusion matrices were calculated during brain-controlled stimulation across all the sessions with intact monkeys during which the treadmill was turned off ( $n = 345$  and  $n = 127$  temporal windows for Q1 and Q2). False-positive detections were rare.



**Extended Data Figure 6 | Modulation of leg locomotor movements during brain-controlled stimulation in intact monkeys.** **a**, Principal component analysis was applied on individual gait cycles extracted from locomotion on a treadmill without stimulation ( $n = 125$  and  $n = 119$  for Q1 and Q2) and during brain-controlled stimulation of the extension ( $n = 33$  and  $n = 54$  gait cycles with stimulation for Q1 and Q2, respectively) or flexion ( $n = 98$  and  $n = 120$  total stimulation steps for Q1 and Q2, respectively) hotspots for the intact monkeys Q1 and Q2. Conventions are the same as in Fig. 3 and Extended Data Fig. 7. This analysis emphasizes the graded modulation of gait parameters when increasing the frequency or amplitude of stimulation for extension and

flexion hotspots independently. **b**, Plots showing relationships between the amplitude or frequency of extensor (magenta) or flexor (cyan) hotspot stimulation and relevant kinematic or muscle activity parameter related to the extension or flexion of the right leg. The upper right plot shows the average foot trajectories during each experimental condition, illustrating the graded modulation of flexion during brain-controlled stimulation of the flexion hotspot. The lower right plot highlights the high degree of leg-specific modulation over the entire extent of tested stimulation parameters. Results were comparable in monkeys Q1 and Q2. \* $P < 0.05$ , \*\* $P < 0.001$ , Wilcoxon rank-sum test. Error bars, s.e.m.

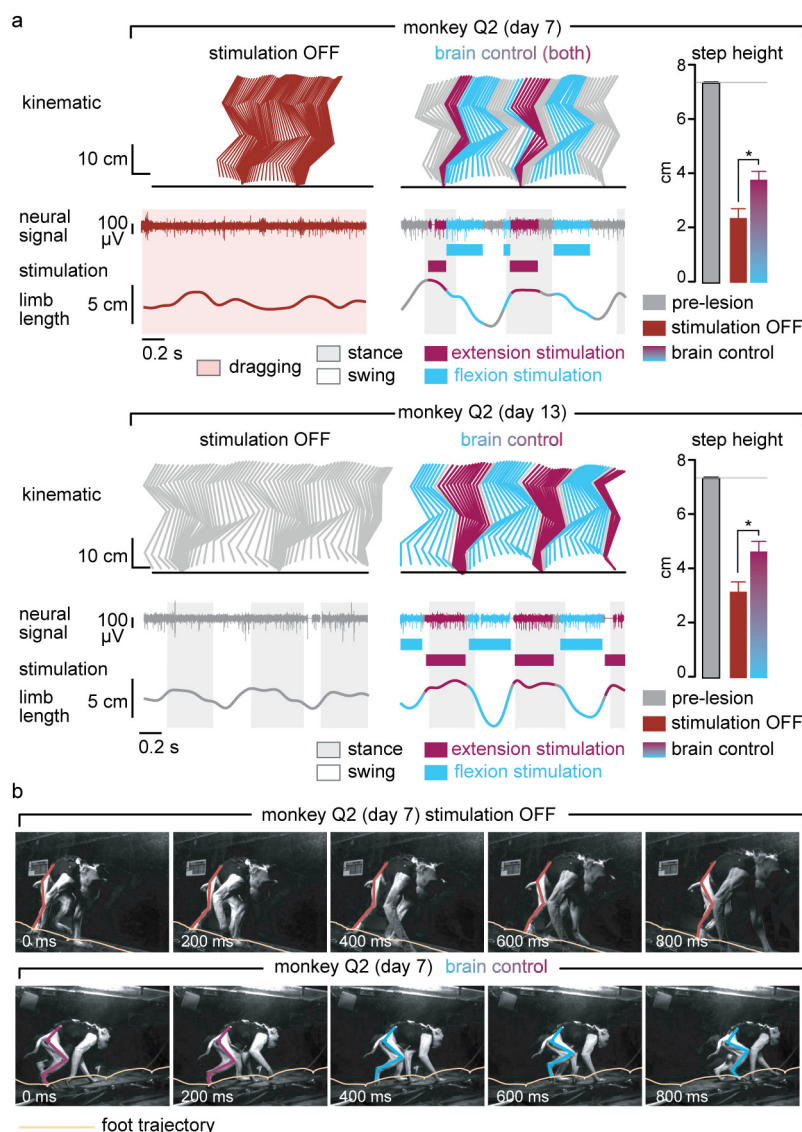




Extended Data Figure 7 | See next page for caption.

**Extended Data Figure 7 | Quantification and reconstruction of the spinal cord lesions and quantification of gait deficits and spontaneous recovery after the spinal cord lesion.** **a**, Scheme illustrating the anatomical experiments to quantify the lesion of the corticospinal tract. The anterograde anatomical tracer BDA was injected into the leg and trunk regions of the left primary motor cortex to label corticospinal tract fibres in the spinal cord, shown in pink in the photographs. The top-right scheme shows the approximate pathways of the dorsal ascending tract and all descending tracts identified from refs 36 and 37. For monkeys Q2 and Q3, the area of maximal damage was reconstructed in two dimensions by identifying the border of the glia scar on evenly spaced sagittal sections spanning the entire dorsoventral extent of the lesioned spinal cord. **b–e**, For each monkey, confocal photographs show longitudinal sections of the lesioned spinal cord at specific dorsoventral levels, as indicated with dotted lines. In each photograph, the following anatomical elements are labelled: astrocytes (GFAP, grey), neural cell bodies (NeuN, cyan) and corticospinal tract axons (BDA, pink). The insets show high-resolution photographs of selected regions (white squares) of the same photograph

that illustrates intact corticospinal tract axons above the injury, axon retraction bulbs directly above the lesion, and the absence of axons below the injury. Monkey Q2 displayed a small subset of spared corticospinal tract axons in the more dorsal aspect of the dorsolateral column. The asterisks indicate the location of the lesion. Overview scale bars, 500  $\mu\text{m}$ . Inset scale bars, 50  $\mu\text{m}$ . **f**, Gait cycles were extracted from locomotion recorded in monkeys Q2 and Q3 pre-lesion, during the first two weeks after the lesion, and at 99 days after the lesion. Analysis was only applied to gait cycles classified as steps, that is, gait cycles classified as limb paralysis or stumbling were not included. The number of analysed gait cycles is directly reported into the figure. The bar plots report the mean Euclidean distance between all the steps on a given recording day without stimulation and steps recorded before the lesion, computed in the entire space of the 26 kinematic parameters. The other bar plots report mean values of relevant kinematic parameters. This analysis illustrates the progressive yet incomplete recovery of locomotion. \*\*\* $P < 0.001$ , Wilcoxon rank-sum test. Error bars, s.e.m.

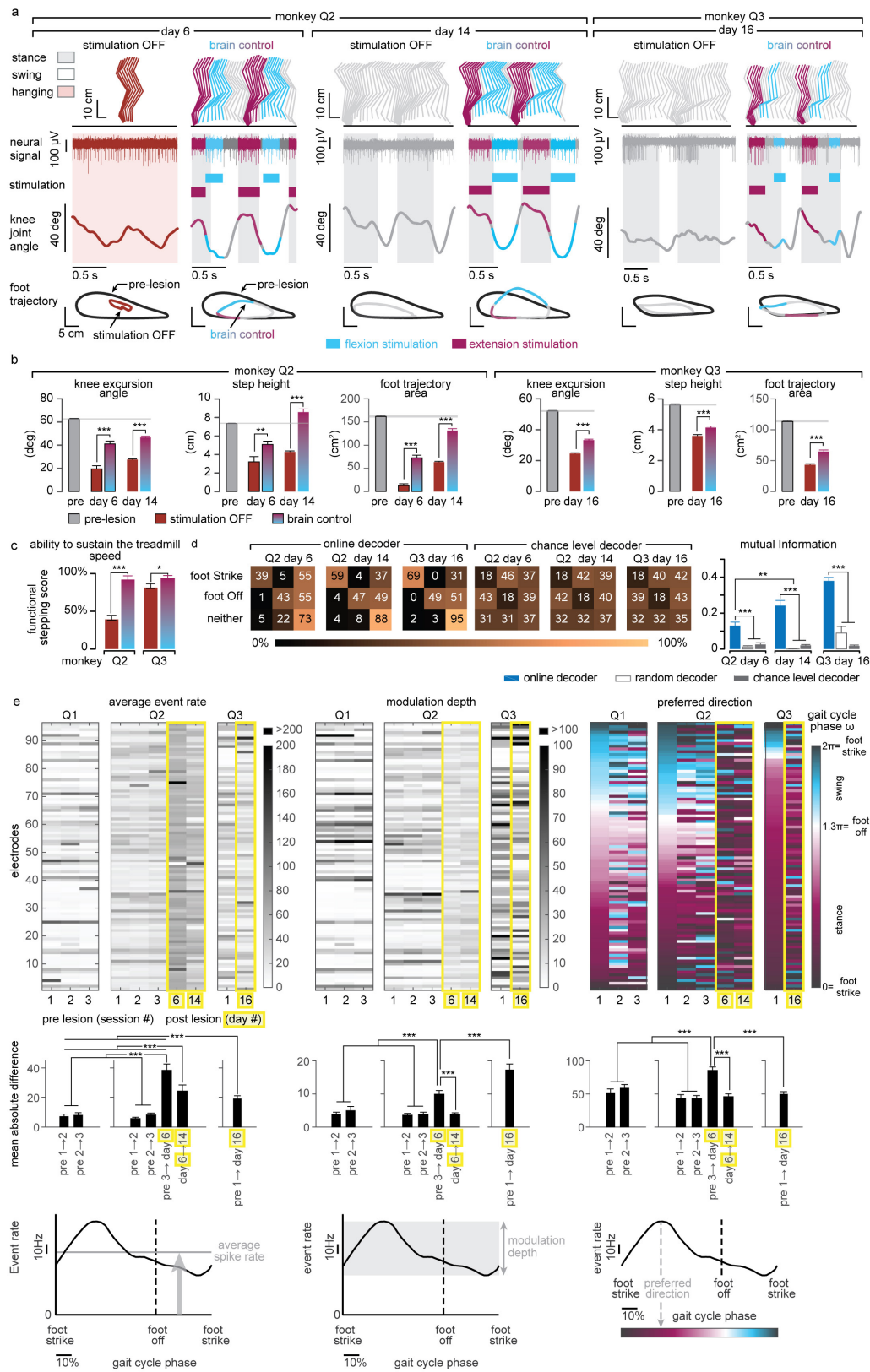


### Extended Data Figure 8 | Restoration of leg locomotor movements during overground locomotion after the spinal cord lesion.

**a**, Representative sequences of locomotion along a straight corridor without stimulation and during brain-controlled stimulation recorded at 7 days and 13 days after the lesion for monkey Q2. Conventions are the same as in Fig. 4. The bar plots report the mean step height of the right (lesioned) leg during swing. \* $P < 0.05$ , Wilcoxon rank-sum test. Error bars, s.e.m. **b**, Snapshots extracted from video recordings showing a

representative sequence of leg movements during one gait cycle without stimulation and during brain-controlled stimulation at 7 days post-injury for monkey Q2. The stick diagrams overlying the right (lesioned) leg and the trajectory of the foot are colour-coded using the same colour scheme as in the stick diagram decomposition in **a**. Without stimulation, the monkey dragged the leg along the ground, whereas brain-controlled stimulation restored weight-bearing locomotion with plantar placement.

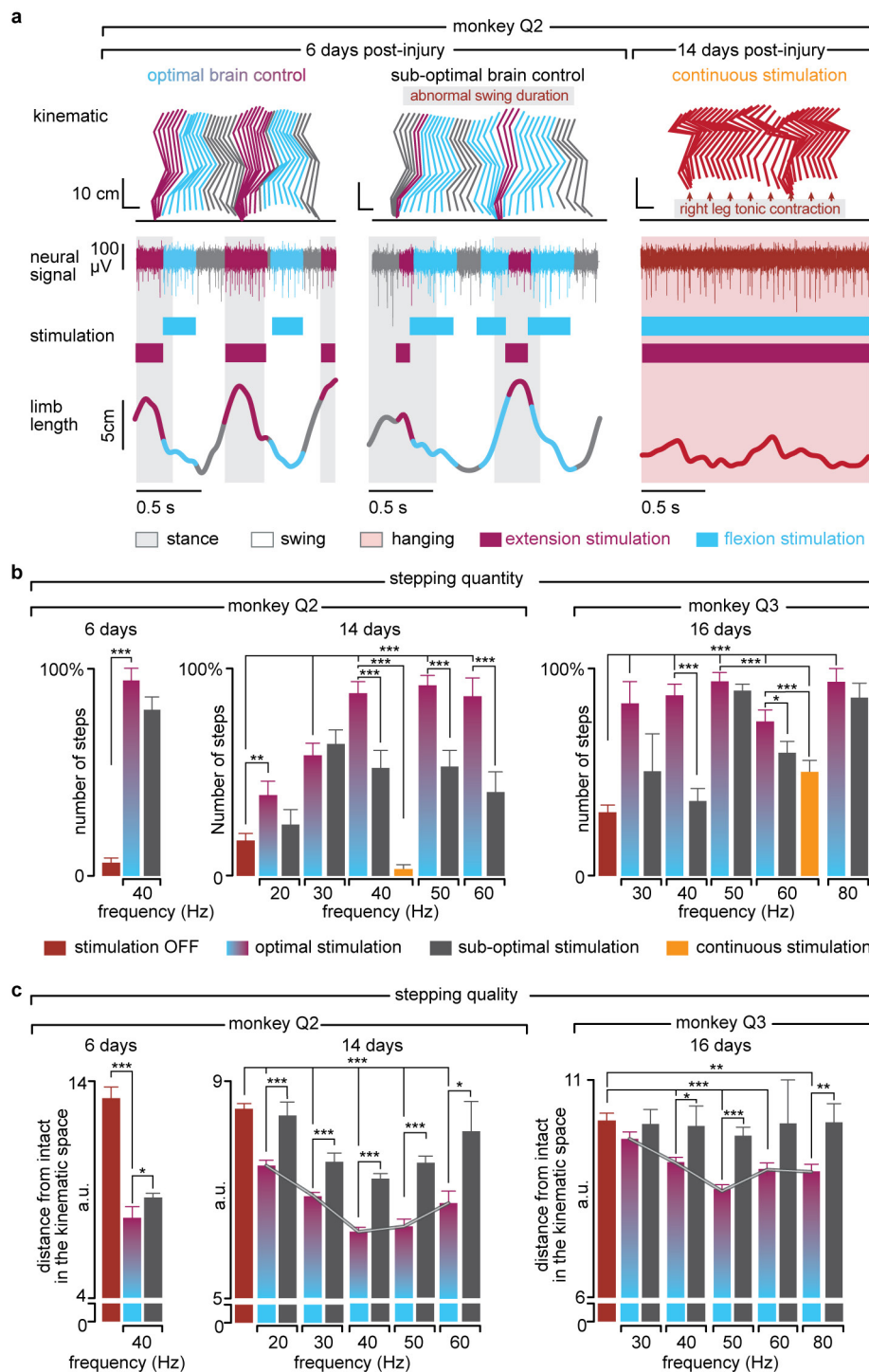




Extended Data Figure 9 | See next page for caption.

**Extended Data Figure 9 | Quantification of gait improvements and decoding accuracy during brain-controlled stimulation after the spinal cord lesion.** **a**, Two successive gait cycles performed during locomotion on a treadmill without stimulation and during brain-controlled stimulation in monkey Q2 at 6 days and 14 days post-lesion, and in monkey Q3 at 16 days post-injury. Conventions are the same as in Fig. 4. In addition, the average foot trajectories calculated over all the recorded gait cycles are displayed for each experimental condition including during pre-lesion locomotion, illustrating the marked improvement of foot movements during brain-controlled stimulation. **b**, Bar plot reporting the mean values of the total excursion of the angle, step height and foot trajectory area for monkeys Q2 and Q3 during locomotion pre-lesion and post-lesion without stimulation and with brain-controlled stimulation. Analysis was only applied to gait cycles classified as steps, that is, gait cycles classified as limb paralysis or stumbling were not included. (For monkey Q2: pre-lesion  $n = 294$ , day 6 post-lesion no stimulation  $n = 6$ , brain control  $n = 12$ , day 14 no stimulation  $n = 39$ , brain control  $n = 93$ . For monkey Q3: pre-lesion  $n = 185$ , day 16 post-injury no stimulation  $n = 98$ , brain control  $n = 31$ .)  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ , Wilcoxon rank-sum test. Error bars, s.e.m. **c**, Bar plots reporting the capacity of the monkeys to sustain walking at the imposed treadmill belt speed. The functional score is computed as the percentage of regular steps in which the animal is able to walk at the treadmill belt speed; that is, the animal does not bump into the back of the treadmill. Gait cycles classified as hops or stumbling were not included.  $*P < 0.05$ ,  $***P < 0.001$ , bootstrap. Error bars, s.e.m. **d**, Decoding

accuracy increases during recovery after the spinal cord lesion. Decoder confusion matrices calculated reporting the accuracy of the real-time decoders and chance level decoder during brain-controlled stimulation for monkey Q2 at day 6 ( $n = 76$  foot-strikes,  $n = 74$  foot-offs) and day 14 post-injury ( $n = 264$  foot-strikes,  $n = 264$  foot-offs) and for monkey Q3 at 16 days post-injury ( $n = 319$  foot-strikes,  $n = 321$  foot-offs). The tolerance window was set at  $\pm 125$  ms. The bar plots report the normalized mutual information calculated for the real-time decoders compared to random decoders and chance level decoders.  $***, **P < 0.001$  and  $P < 0.01$ , respectively. Bootstrap. Error bars, s.e.m. **e**, Top panels, from left to right: mean event rate, modulation depth and preferred direction for the neuronal spiking signal recorded obtained by regressing spike rates against the phase of the gait cycle for monkeys Q1 to Q3. Preferred direction was defined as the angle for which the fitted tuning function was at a maximum. Bottom panels, from left to right: mean absolute single-electrode difference for mean event rates, modulation depths and preferred directions between two consecutive sessions shown in the top panels. Analysis shows substantial changes both before and after the spinal-cord lesion. Nevertheless, the rate of change between the last pre-lesion and the first post-lesion sessions was substantially higher than between any two other session pairs, thus indicating increased level of plasticity following spinal cord lesion.  $**P < 0.01$ ,  $***P < 0.001$ , Monte Carlo, Wilcoxon rank-sum test, signed Wilcoxon rank-sum test and bootstrap. Error bars, s.e.m.



**Extended Data Figure 10 | The temporal structure and features of stimulation determine the quantity and quality of steps. a,** Two successive gait cycles performed during locomotion on a treadmill during brain-controlled stimulation with optimal and suboptimal temporal structures at 6 days post-injury, and during continuous stimulation at 14 days post-lesion for monkey Q2. The gait cycles were classified as suboptimal temporal structures when stimulation occurred outside a  $\pm 125$  ms tolerance window. Conventions are the same as in Fig. 4. **b,** Bar plots reporting the quantity of steps calculated during locomotion with optimal and suboptimal temporal structures over a range of stimulation frequencies and during continuous stimulation for monkeys Q2 and Q3 at 6 days, 14 days and 16 days post-injury.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ , bootstrap. Error bars, s.e.m. **c,** Bar plots reporting the

quality of stepping for the same conditions as in **b**. The quality of stepping was measured as the mean Euclidean distance between pre-lesion and post-lesion gait cycles calculated in the kinematic space defined by the 26 gait parameters, as reported in Fig. 4. These results show that the optimal temporal structure leads to an increased number of steps and improved quality of stepping compared to sub-optimal temporal structures. Moreover, brain-controlled stimulation with both optimal and suboptimal temporal structure promoted markedly improved locomotor performance compared to continuous stimulation delivered with the same stimulation features. These results also highlight the ability to optimize locomotor performance when tuning the stimulation frequency.  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ , Wilcoxon rank-sum test. Error bars, s.e.m.



## LETTER

doi:10.1038/nature19845

**A basal ganglia circuit for evaluating action outcomes**

Marcus Stephenson-Jones, Kai Yu, Sandra Ahrens, Jason M. Tucciarone, Aile N. van Huijstee, Luis A. Mejia, Mario A. Penzo, Lung-Hao Tai, Linda Wilbrecht, Bo Li

This is a PDF file of a peer-reviewed paper that has been accepted for publication. Although unedited, the content has been subjected to preliminary formatting. *Nature* is providing this early version of the typeset paper as a service to our customers. The text and figures will undergo copyediting and a proof review before the paper is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply.

Cite this article as: Stephenson-Jones, M. *et al.* A basal ganglia circuit for evaluating action outcomes. *Nature* <http://dx.doi.org/10.1038/nature19845> (2016).

# A basal ganglia circuit for evaluating action outcomes

Marcus Stephenson-Jones<sup>1</sup>, Kai Yu<sup>1</sup>, Sandra Ahrens<sup>1</sup>, Jason M. Tucciarone<sup>1</sup>, Aile N. van Huijstee<sup>1</sup>, Luis A. Mejia<sup>1</sup>, Mario A. Penzo<sup>1</sup>, Lung-Hao Tai<sup>2</sup>, Linda Wilbrecht<sup>2,3</sup>, Bo Li<sup>1</sup>

**The basal ganglia, a group of subcortical nuclei, play a crucial role in decision making by selecting actions and evaluating their outcomes<sup>1,2</sup>. While much is known about the function of the basal ganglia circuitry in selection<sup>1,3,4</sup>, how these nuclei contribute to outcome evaluation is less clear. Here we show that neurons in the habenula-projecting globus pallidus (GPh) are essential for evaluating action outcomes and are regulated by a specific set of inputs from the basal ganglia. We found in a classical conditioning task that individual mouse GPh neurons bidirectionally encode whether an outcome is better or worse than expected. Mimicking these evaluation signals with optogenetic inhibition or excitation is sufficient to reinforce or discourage actions in a decision making task. Moreover, cell-type-specific synaptic manipulations revealed that the inhibitory and excitatory inputs to the GPh are necessary for mice to appropriately evaluate positive and negative feedback, respectively. Finally, using rabies virus-assisted monosynaptic tracing<sup>5</sup>, we discovered that the GPh is embedded in a basal ganglia circuit wherein it receives inhibitory input from both striosomal and matrix compartments of the striatum, and excitatory input from the “limbic” regions of the subthalamic nucleus (STN). Our results provide the first direct evidence that information about the selection and evaluation of actions is channelled through distinct sets of basal ganglia circuits, with the GPh representing a key locus where information of opposing valence is integrated to determine whether action outcomes are better or worse than expected.**

The GPh, a phylogenetically conserved non-motor output of the basal ganglia<sup>6–8</sup>, excites the lateral habenula (LHb) that, in turn, drives inhibition onto dopamine neurons when an outcome is worse than expected<sup>8–11</sup>. GPh neurons may thus play a key role in evaluating action outcomes by providing a source of “prediction error (PE)” to the reward system, to drive reinforcement learning. In order to test this hypothesis, we first verified that we could selectively target GPh neurons in the entopeduncular nucleus (EP), the rodent homolog of the primate globus pallidus interna (GPi) where GPh neurons are located<sup>7,8,12,13</sup>, on the basis of their expression of vesicular glutamate transporter-2 (Vglut2) and the neuropeptide somatostatin (Som)<sup>6,8,14</sup>, and that GPh neurons project exclusively to the LHb<sup>13</sup> (Extended Data Fig. 1).

To examine the function of GPh neurons in relation to outcome evaluation, we modified a classical conditioning task designed for studying value coding in dopamine neurons<sup>15</sup>. Here, a unique auditory conditioned stimulus (CS) predicted the delivery of one of five unconditioned stimuli (US): water rewards (1 and 5  $\mu$ l), nothing, or air puffs to the face (100 and 500 ms). As training progressed, mice began licking or blinking in response to the reward or punishment-predicting cues, respectively. The lick rate and blinking occurrence were significantly higher for cues that predicted large rewards and punishments than for cues predicting small rewards and punishments (Fig. 1a, b), indicating that mice had learnt the CS-US associations.

We recorded the activity of EP neurons in *Vglut2-Cre;Ai35* mice, in which Vglut2<sup>+</sup> GPh neurons could be optogenetically tagged with the light-sensitive proton pump archaerhodopsin (Arch) (see Methods) (Extended Data Fig. 2a–c), while they performed this task. Hierarchical clustering revealed that all the tagged neurons belonged to one class of neurons, which we classified as putative GPh neurons (Extended Data Fig. 2d–g). In contrast, neurons in two other functional clusters were never optogenetically tagged, and showed an activity profile resembling that of the classic GABAergic movement-related neurons found in the GPi<sup>16</sup> (Extended Data Fig. 2d–j).

Putative GPh neurons were phasically excited by both punishment-predicting tones and punishments (Fig. 1c, Extended Data Fig. 2h). In a portion of these neurons, the magnitude of single-neuron tone responses was greater when the tone predicted a larger punishment (Fig. 1c,e). In addition, as in primates<sup>7,17</sup>, GPh neurons were phasically inhibited by both reward-predicting tones and rewards (Fig. 1c, Extended Data Fig. 2d–f); the inhibition was greater when the tone predicted a larger reward (Fig. 1c,d). The average magnitude of the CS responses in these neurons was graded, reflecting the expected magnitude of reward or punishment (Fig. 1f–i). We conclude that individual GPh neurons bidirectionally encode the expected value of an action as well as the value of its outcome.

To determine how these expectation and outcome signals develop, we recorded GPh neurons over the course of the behavioural training (Extended Data Fig. 3a). Prior to training, GPh neurons showed no response to the CS (Extended Data Fig. 3b, c). During training, these neurons rapidly acquired a CS response, which was initially smaller but became gradually larger than the US response as training proceeded (Fig. 2a, b, e, f; Extended Data Fig. 3d, e). Indeed, once the animals had fully learnt the task, and thus could readily predict the US, the response to the US (reward or punishment) was markedly suppressed or even absent in the majority of neurons (Fig. 2a, b, e, f; Extended Data Fig. 3d, e). However, unexpected delivery of the US still evoked a response, even in neurons that no longer responded to the US when it was predicted (Fig. 2a, e, Extended Data Fig. 3d, e). This reduction in US response in GPh neurons is consistent with encoding of PE, a function well described for dopamine neurons<sup>15,18</sup>.

Another hallmark of PE coding is that neurons respond when an expected outcome is omitted<sup>15,19</sup>. To test if GPh neurons have such property, we omitted an expected US in 10% of the large reward or punishment trials. When an expected punishment was omitted, putative GPh neurons displayed a decrease in firing either compared to when the punishment was delivered (Fig. 2c, Extended Data Fig. 3f, g), or compared with baseline (Fig. 2d). In contrast, upon omission of an expected reward, putative GPh neurons showed an increase in firing rate relative to delivery of the reward (Fig. 2g, Extended Data Fig. 3h, i), and relative to baseline (Fig. 2h). Together, these results demonstrate that GPh neurons encode reward and punishment PEs, bidirectionally signaling when an outcome is better or worse than expected.

<sup>1</sup>Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. <sup>2</sup>Department of Psychology, University of California Berkeley, Berkeley, CA 94720, USA. <sup>3</sup>Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA 94720, USA.

The observed bidirectional responses to reward and punishment suggest that inhibition or excitation of the GPh may influence behaviour. While excitation of the GPh is aversive<sup>8</sup>, inhibition of it may be rewarding. To test this, we introduced Arch selectively into GPh neurons of *Vglut2-Cre* mice and optogenetically inhibited these neurons when the mice performed a real-time place preference task. Inhibition of either the somata of GPh neurons (Extended Data Fig. 4a–c) or their axon terminals in the LHb (Extended Data Fig. 4d–f) induced a preference in these mice for the inhibition-paired chamber (Extended Data Fig. 5a–i). Mice would also actively work to receive inhibition of the GPh (Extended Data Fig. 5j). These effects were not specific to the *Vglut2-Cre* mice, as optogenetic inhibition or excitation (using channelrhodopsin (ChR2), the light-gated cation channel) of the GPh, targeted in wild-type mice with a retrograde canine adenovirus (CAV2-Cre)<sup>20</sup>, also induced real-time place preference or aversion, respectively (Extended Data Fig. 4g–l, 5k–m). None of our optogenetic manipulations had an effect on the velocity or distance of movement (Extended Data Fig. 5n,o). These data show that excitation and inhibition of the GPh have opposing motivational valence, with the former being aversive and the latter rewarding.

GPh neurons' response properties and their influence on behaviour point to the possibility that they could play a fundamental role in evaluating action outcomes. To test this possibility, we trained *Vglut2-Cre* mice in a probabilistic switching task, where animals had to rely on the evaluation of previous choice outcomes to update their future decisions (Extended Data Fig. 6a–f). Previous studies have shown that mice adopt a “win-stay, lose-switch” strategy in this task that can be best described by a multivariate logistic regression model<sup>21</sup> (Extended Data Fig. 6a–f; also see Methods).

We used optogenetics to specifically inhibit or activate GPh neurons (Extended Data Fig. 4a–c, m–o) at the moment of outcome evaluation in 10% of the trials in this task. Inhibition of GPh neurons at the time that a mouse nose poked in a water port significantly biased the mouse to return to the same port on the subsequent trial (Fig. 3a, b, Supplementary Video 1). Conversely, activation of the GPh–LHb pathway when a mouse nose poked in a water port significantly promoted the mouse to switch to the alternative port on the subsequent trial (Fig. 3d, e, Supplementary Video 2). In both cases, the probability of repeating the same choice was dependent on both the optogenetic manipulation and the previous reward history. Optogenetic inhibition or activation of GPh neurons shifted the sigmoidal decision curve along the x-axis (Fig. 3b, e), indicating these manipulations mimicked a fixed increase or decrease, respectively, in the value of a chosen action (Extended Data Fig. 6g–i; also see Methods). Alternate models did not provide a better explanation for our data (Supplementary table 1). In control experiments, light illumination of GPh neurons expressing eYFP produced no effect on choice behaviour (Fig. 3c, f, Extended Data Fig. 6g,h). These results indicate that bidirectional changes in GPh activity are sufficient to bias outcome evaluation thereby reinforcing or discouraging particular actions.

To examine whether the GPh also contributes to action selection, we optogenetically inhibited or activated the GPh in the same mice, but at the time when they nose poked the central port to initiate a trial, a time-frame that may coincide or overlap with the moment of action selection<sup>21</sup> (Extended Data Fig. 7). Neither of these manipulations produced an effect on choice (Extended Data Fig. 7a–h). In addition, these manipulations did not appear to influence ongoing behaviour (Extended Data Fig. 7i–l). Thus, activation or inhibition of the GPh influences the evaluation, but not selection, of actions.

While these optogenetic experiments demonstrate that bidirectional changes in GPh activity are sufficient to reinforce or discourage actions, they do not determine whether endogenous GPh activity is required for such function. We reasoned that if the excitatory and inhibitory responses of GPh neurons to action outcomes are essential in providing negative and positive feedback, respectively, then reducing the excitatory or inhibitory inputs onto these neurons should accordingly impair

choice behaviour. To test this hypothesis, we first selectively weakened glutamatergic synapses onto GPh neurons by expressing in these neurons the C-terminal tail of AMPA receptor (AMPA) subunit GluA4 (GluA4-ct) (Extended Data Fig. 8a,e). The GluA4-ct inhibits excitatory synaptic transmission by blocking AMPAR synaptic trafficking<sup>22</sup> (Extended Data Fig. 8b–d). Interestingly, mice expressing the GluA4-ct in the GPh (GPh<sup>GluA4-ct</sup>) were significantly less likely to switch their choice following an unrewarded outcome (lose-switch) (Fig. 4a, b), and were slower to reverse their choice when the reward contingencies were switched (Fig. 4c). On the other hand, GPh<sup>GluA4-ct</sup> mice had no change in the ability to repeat the same choice following a rewarded outcome (win-stay) (Extended Data Fig. 8f). Logistic regression analysis revealed that GPh<sup>GluA4-ct</sup> mice had a substantial reduction in their negative regression coefficients, the weighted contribution that past unrewarded trials had on current choice (Fig. 4d,e, Extended Data Fig. 8g), but had no change in their positive regression coefficients (Extended Data Fig. 8h, i). These results indicate that reducing the glutamatergic transmission onto the GPh impairs negative feedback by selectively diminishing the impact of unrewarded outcomes on future decisions.

Next, we selectively weakened GABAergic synapses onto GPh neurons. To this end, we injected the EP of *Som-Flp; Gabrg2<sup>fllox</sup>* mice<sup>23</sup> with a virus expressing Cre in a Flp-dependent manner<sup>22</sup> (Extended Data Fig. 8j, n). In these mice only GPh neurons within the EP could express Cre and thus have the  $\gamma 2$  subunit of GABA<sub>A</sub> receptor ablated (GPh <sup>$\gamma 2$ -KO</sup>). As expected, this approach led to a significant reduction of GABA<sub>A</sub>-mediated synaptic transmission onto GPh neurons (Extended Data Fig. 8k–m). Compared with control mice, GPh <sup>$\gamma 2$ -KO</sup> mice were less persistent in response to positive feedback (Fig. 4f), showed a substantial reduction in win-stay (Fig. 4g) but no change in lose-switch behaviour (Extended Data Fig. 8o), and were faster to reverse their choice when the reward contingencies were switched (Fig. 4h), presumably due to decreased sensitivity to reward. Logistic regression analysis revealed that the positive regression coefficients – the weighted contributions that past rewarded trials had on current choice – were reduced in GPh <sup>$\gamma 2$ -KO</sup> mice (Fig. 4i, j, Extended Data Fig. 8p). In contrast, there was no overall change in the negative regression coefficients in GPh <sup>$\gamma 2$ -KO</sup> mice (Extended Data Fig. 8q, r). These results indicate that reducing the inhibitory input onto the GPh impairs positive feedback by selectively reducing the impact past rewarded outcomes had on future choice.

To identify the circuits upstream of the GPh that may provide the reward and punishment information, we used a modified rabies virus system to trace the monosynaptic inputs onto GPh neurons<sup>5</sup> (Extended Data Fig. 9a–c). We found that, like the canonical basal ganglia output nuclei GPi/SNr (substantia nigra pars reticulata)<sup>4</sup>, the GPh received inputs directly from the striatum (Fig. 5a, Extended Data Fig. 9d). However, unlike the GPi/SNr<sup>12,24</sup>, a large proportion of the inputs to the GPh arose from the striosomal compartment of the striatum (Fig. 5b, c). These striatal inputs could drive monosynaptic GABAergic responses in the GPh (Fig. 5d, e). These data indicate that reward related evaluation signals in the GPh may at least in part arise from subsets of neurons in both the striosomal and matrix compartments of the striatum.

In addition to this direct projection from the striatum, we found that the GPh was also regulated by distinct nuclei associated with the “indirect pathway”. In contrast to the GPi, which receives excitatory input from the core of the STN<sup>25</sup> and input from parvalbumin positive GPe neurons<sup>26</sup>, the GPh received input from the subthalamic cells located in the “limbic” region of the STN<sup>27</sup>, on the medial border of this nucleus and in the surrounding paraventricular nucleus (pSTN), and GPe input from mainly parvalbumin negative neurons (Fig. 5a,f, Extended Data Fig. 9e–i). This medial “limbic” STN could drive monosynaptic excitation in GPh neurons (Fig. 5g,h), and may thus provide these neurons with the negative valence information<sup>28</sup>. These results indicate that the GPh is embedded in a basal ganglia circuit that is



intermingled with but distinct from the circuitry that regulates the GPi/SNr (Extended Data Fig. 9j).

Together, our results demonstrate that the GPh is a key locus where information of opposing valence is integrated, from a subset of basal ganglia circuits, to determine if an action is better or worse than expected. The outcome evaluation function of the GPh is likely mediated through bidirectional control of dopamine neurons, in which PE coding is critical for reinforcement learning. The GPh is well placed to bidirectionally influence dopaminergic activity as it provides tonic excitatory input to the LHB<sup>6,8</sup>, which in turn regulates dopamine neurons<sup>7,9,11</sup> disynaptically via the GABAergic rostromedial tegmental nucleus<sup>7,29</sup>. Indeed, bidirectional changes in LHB firing have opposing effects on dopamine cell firing<sup>7,9,11</sup>, and lesions of the LHB disrupt negative and impair positive reward PE coding in dopaminergic neurons<sup>30</sup>. We propose that an increase in GPh activity when an outcome is worse than expected increases the excitatory drive onto the LHB to inhibit dopamine neurons and discourage actions, whereas decreases in GPh activity when an outcome is better than expected remove the tonic excitation of the LHB to increase dopaminergic activity and reinforce actions (Extended Data Fig. 10).

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 11 October 2015; accepted 14 September 2016.**

**Published online 21 September 2016.**

- Nelson, A. B. & Kreitzer, A. C. Reassessing models of basal ganglia function and dysfunction. *Annu Rev Neurosci* **37**, 117–135, doi:10.1146/annurev-neuro-071013-013916 (2014).
- Amemori, K., Gibb, L. G. & Graybiel, A. M. Shifting responsibility: the importance of striatal modularity to reinforcement learning in uncertain environments. *Front Hum Neurosci* **5**, 47, doi:10.3389/fnhum.2011.00047 (2011).
- Hikosaka, O. Basal ganglia mechanisms of reward-oriented eye movement. *Ann N Y Acad Sci* **1104**, 229–249, doi:10.1196/annals.1390.012 (2007).
- Alexander, G. E. & Crutcher, M. D. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci* **13**, 266–271 (1990).
- Callaway, E. M. & Luo, L. Monosynaptic Circuit Tracing with Glycoprotein-Deletable Rabies Viruses. *J Neurosci* **35**, 8979–8985, doi:10.1523/JNEUROSCI.0409-15.2015 (2015).
- Stephenson-Jones, M., Kardamakis, A. A., Robertson, B. & Grillner, S. Independent circuits in the basal ganglia for the evaluation and selection of actions. *Proc Natl Acad Sci U S A* **110**, E3670–E3679, doi:10.1073/pnas.1314815110 (2013).
- Hong, S. & Hikosaka, O. The globus pallidus sends reward-related signals to the lateral habenula. *Neuron* **60**, 720–729, doi:10.1016/j.neuron.2008.09.035 (2008).
- Shabel, S. J., Proulx, C. D., Trias, A., Murphy, R. T. & Malinow, R. Input to the lateral habenula from the basal ganglia is excitatory, aversive, and suppressed by serotonin. *Neuron* **74**, 475–481, doi:10.1016/j.neuron.2012.02.037 (2012).
- Matsumoto, M. & Hikosaka, O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**, 1111–1115 (2007).
- Hong, S., Zhou, T. C., Smith, M., Saleem, K. S. & Hikosaka, O. Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *J Neurosci* **31**, 11457–11471, doi:10.1523/JNEUROSCI.1384-11.2011 (2011).
- Stamatakis, A. M. & Stuber, G. D. Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. *Nat Neurosci* **15**, 1105–1107, doi:10.1038/nn.3145 (2012).
- Rajakumar, N., Elisevich, K. & Flumerfelt, B. A. Compartmental origin of the striato-entopeduncular projection in the rat. *J Comp Neurol* **331**, 286–296, doi:10.1002/cne.903310210 (1993).
- Parent, M., Levesque, M. & Parent, A. Two types of projection neurons in the internal pallidum of primates: single-axon tracing and three-dimensional reconstruction. *J Comp Neurol* **439**, 162–175 (2001).
- Vincent, S. R. & Brown, J. C. Somatostatin immunoreactivity in the entopeduncular projection to the lateral habenula in the rat. *Neurosci Lett* **68**, 160–164 (1986).
- Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85–88, doi:10.1038/nature10754 (2012).
- DeLong, M. R., Crutcher, M. D. & Georgopoulos, A. P. Primate globus pallidus and subthalamic nucleus: functional organization. *J Neurophysiol* **53**, 530–543 (1985).
- Bromberg-Martin, E. S., Matsumoto, M., Hong, S. & Hikosaka, O. A pallidum-habenula-dopamine pathway signals inferred stimulus values. *J Neurophysiol* **104**, 1068–1076, doi:10.1152/jn.00158.2010 (2010).
- Pan, W. X., Schmidt, R., Wickens, J. R. & Hyland, B. I. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J Neurosci* **25**, 6235–6242, doi:10.1523/JNEUROSCI.1478-05.2005 (2005).
- Schultz, W. Dopamine reward prediction-error signalling: a two-component response. *Nat Rev Neurosci* **17**, 183–195, doi:10.1038/nrn.2015.26 (2016).
- Bru, T., Salinas, S. & Kremer, E. J. An update on canine adenovirus type 2 and its vectors. *Viruses* **2**, 2134–2153, doi:10.3390/v2092134 (2010).
- Tai, L. H., Lee, A. M., Benavidez, N., Bonci, A. & Wilbrecht, L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci* **15**, 1281–1289, doi:10.1038/nn.3188 (2012).
- Ahrens, S. et al. ErbB4 regulation of a thalamic reticular nucleus circuit for sensory selection. *Nat Neurosci* **18**, 104–111, doi:10.1038/nn.3897 (2015).
- Wulff, P. et al. From synapse to behavior: rapid modulation of defined neuronal types with engineered GABA<sub>A</sub> receptors. *Nat Neurosci* **10**, 923–929, doi:10.1038/nn1927 (2007).
- Fujiyama, F. et al. Exclusive and common targets of neostriatofugal projections of rat striosome neurons: a single neuron-tracing study using a viral vector. *Eur J Neurosci* **33**, 668–677, doi:10.1111/j.1460-9568.2010.07564.x (2011).
- Kita, H. & Kitai, S. T. Efferent projections of the subthalamic nucleus in the rat: light and electron microscopic analysis with the PHA-L method. *J Comp Neurol* **260**, 435–452, doi:10.1002/cne.902600309 (1987).
- Mastro, K. J., Bouchard, R. S., Holt, H. A. & Gittis, A. H. Transgenic mouse lines subdivide external segment of the globus pallidus (GPe) neurons and reveal distinct GPe output pathways. *J Neurosci* **34**, 2087–2099, doi:10.1523/JNEUROSCI.4646-13.2014 (2014).
- Hamani, C., Saint-Cyr, J. A., Fraser, J., Kaplitt, M. & Lozano, A. M. The subthalamic nucleus in the context of movement disorders. *Brain* **127**, 4–20, doi:10.1093/brain/awh029 (2004).
- Breyse, E., Pelloux, Y. & Baunez, C. The Good and Bad Differentially Encoded within the Subthalamic Nucleus in Rats (1,2,3). *eNeuro* **2**, doi:10.1523/ENEURO.0014-15.2015 (2015).
- Jhou, T. C., Fields, H. L., Baxter, M. G., Saper, C. B. & Holland, P. C. The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* **61**, 786–800, doi:S0896-6273(09)00121-4 [pii] 10.1016/j.neuron.2009.02.001 (2009).
- Tian, J. & Uchida, N. Habenula Lesions Reveal that Multiple Mechanisms Underlie Dopamine Prediction Errors. *Neuron* **87**, 1304–1316, doi:10.1016/j.neuron.2015.08.028 (2015).

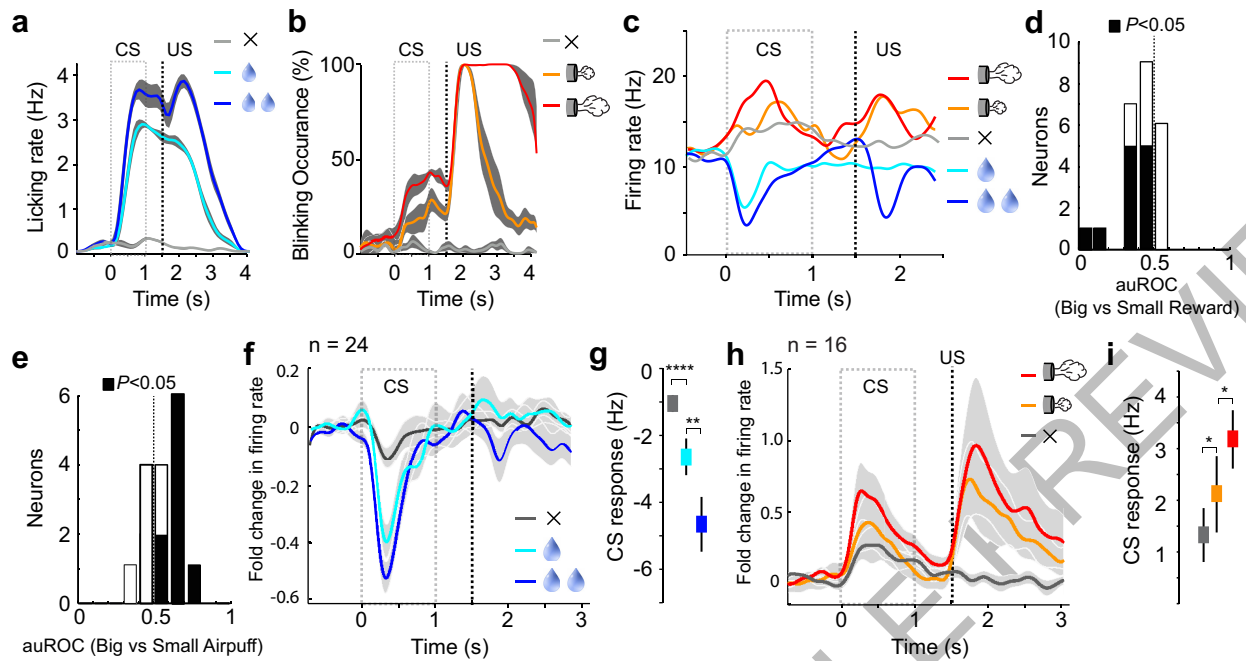
**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank Alissa Cutrone, Danxun Li, and Ga-Ram Hwang for technical assistance, Drs. Vinod Rao and Naoshige Uchida for sharing the Matlab code for the ROC and clustering analysis, Drs. Stephen D. Shea and Saya H. Ebbesen for critical reading of the manuscript, Dr. Z. Josh Huang for providing mouse strains, and members of the Li laboratory for helpful discussions. This work was supported by grants from the National Institutes of Health (NIH) (R01MH108924 to B.L.), the Dana Foundation (to B.L.), NARSAD (to B.L., M.S. and S.A.), Louis Feil Trust (to B.L.), the Stanley Family Foundation (to B.L.), Simons Foundation (to B.L.), Wadecroft Foundation (to B.L.), and an EMBO Long-Term Fellowship Award (to M.S.).

**Author Contributions** M.S. and B.L. designed the study. M.S. conducted experiments and analyzed the data. K.Y. assisted with analysis and implementation of the *in vivo* recording experiments. S.A. and M.P. performed the patch clamp recording experiments. A.N.H. assisted with behavioral training and tracing experiments. J.T. designed and produced the starter virus for the rabies tracing and assisted all the rabies tracing experiments. L.M. assisted with the *in vitro* electrophysiology experiments. L.H.T. and L.W. assisted with the analysis of behavior in the probabilistic switching task. M.S. and B.L. wrote the paper.

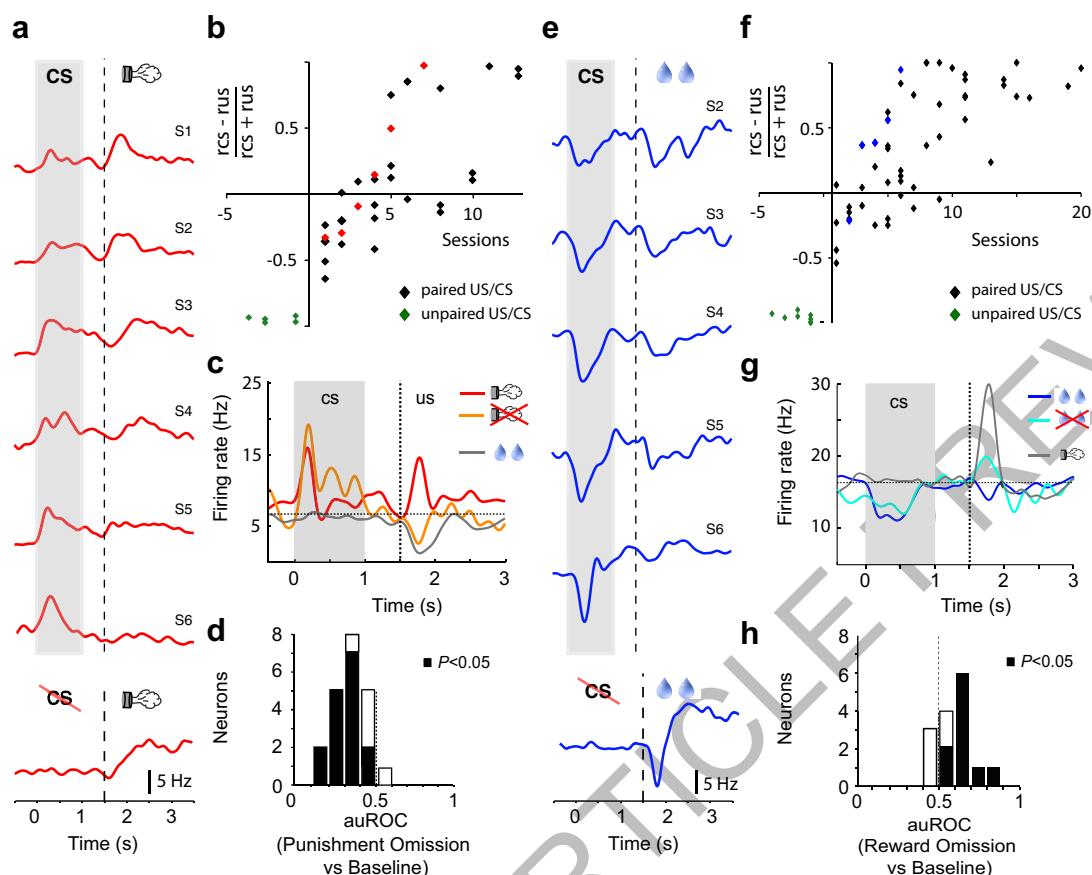
**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.S. (mstephen@cshl.edu) or B.L. (bli@cshl.edu).

**Reviewer Information** Nature thanks G. Stuber and the other anonymous reviewer(s) for their contribution to the peer review of this work.



**Figure 1 | GPh neurons bidirectionally integrate reward and punishment related information.** **a & b.** Licking (a) and blinking (b) behaviour from a representative experimental session. The dashed boxed area and dashed line indicate the time of CS and US delivery, respectively. Licking rate ( $n = 30$  sessions from 7 mice,  $F(2,18) = 41.59$ ,  $P < 0.0001$ ,  $P < 0.05$  for all comparisons) and blinking occurrence ( $n = 32$  sessions from 4 mice,  $F(2,9) = 33.13$ ,  $P < 0.001$ ,  $P < 0.05$  for all comparisons) during the delay between CS and US in recording sessions were compared across different US magnitudes with one-way ANOVA followed by Tukey's test. **c.** Responses of an example putative GPh neuron, shown as spike density functions. **d & e.** auROC (area under the Receiver Operating Characteristic) analysis of differences in firing rate between big and

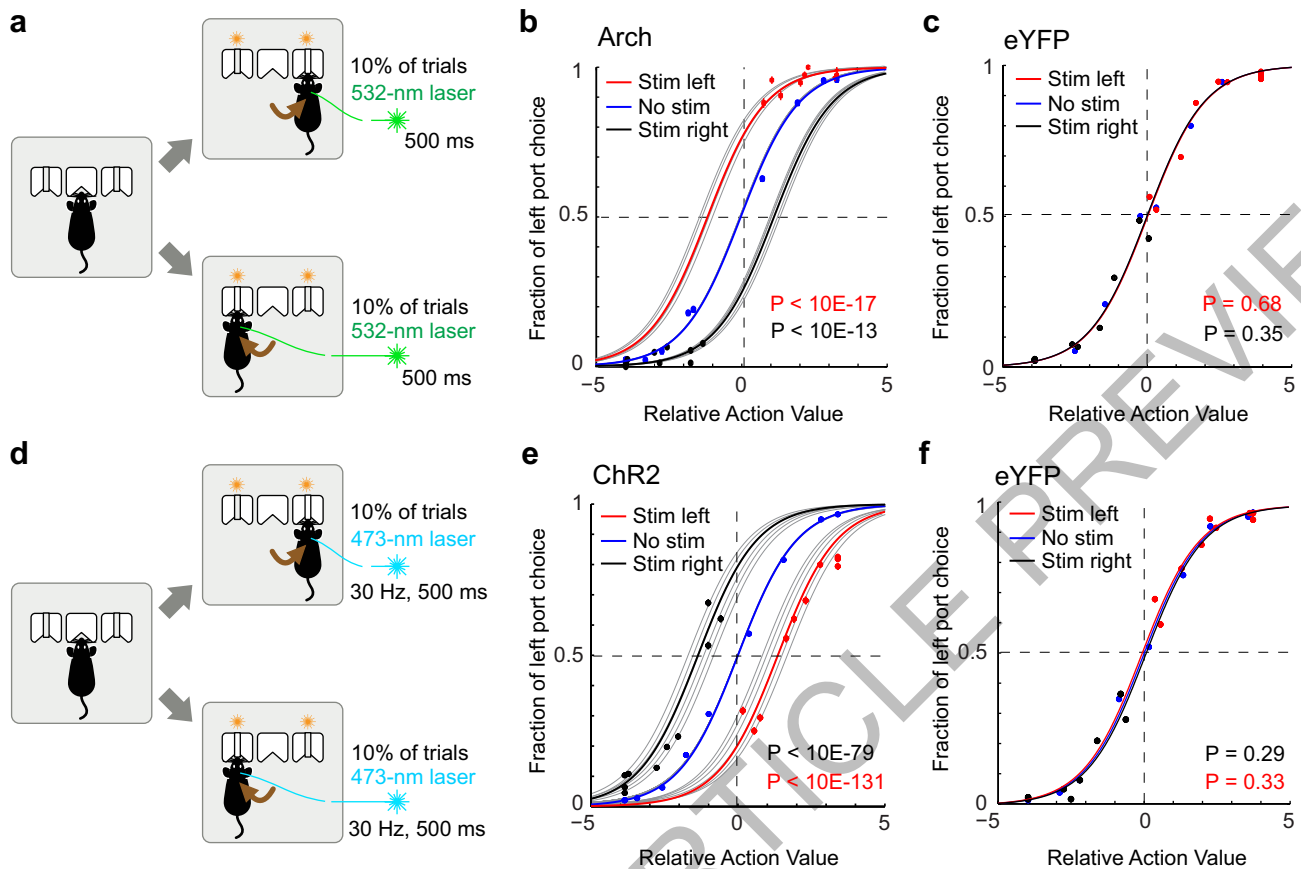
small reward trials (d), or between big and small punishment trials (e), during the peak response to the CS presentation (180–480 ms). Filled bars,  $P < 0.05$ ,  $t$  test. **f.** Average response of putative GPh neurons to reward. **g.** Firing rate change during CS predicting reward of different amplitudes (Big vs. Small reward,  $z = -3.2$ ,  $**P < 0.01$ ; Small vs. No reward,  $z = -4.11$ ,  $***P < 0.0001$ ; Wilcoxon signed-rank test). **h.** Average response of putative GPh neurons to punishment. **i.** Firing rate change during CS predicting punishment of different durations (Big vs. Small punishment,  $z = 2.27$ ,  $*P < 0.05$ ; Small vs. No punishment,  $z = 2.06$ ,  $*P < 0.05$ , Wilcoxon signed-rank test). Data are represented as mean  $\pm$  s.e.m. in **a, b, f–i.**



**Figure 2 | GPh responses to unconditioned stimuli are modulated by expectation.** **a**, CS and US (airpuff) responses of an example putative GPh neuron tracked over multiple sessions. Session-by-session waveform correlations for this individual unit were  $>0.96$ . **b**, CS-US (airpuff) response index for 36 putative GPh neurons (6 mice) across different stages of training (red dots represent values of the sample neuron in **a**) ( $r^2 = 0.56$ ,  $P < 0.0001$  by a linear regression). **c**, Responses of an example GPh neuron to an expected airpuff (red), an unexpectedly omitted airpuff (orange), or an unsignalled reward (grey). **d**, auROC analysis of differences in firing rate between baseline and US presentation time (1.7–1.9 s) in airpuff omission trials ( $n = 21$  neurons from 6 mice). Filled bars,  $P < 0.05$ ,

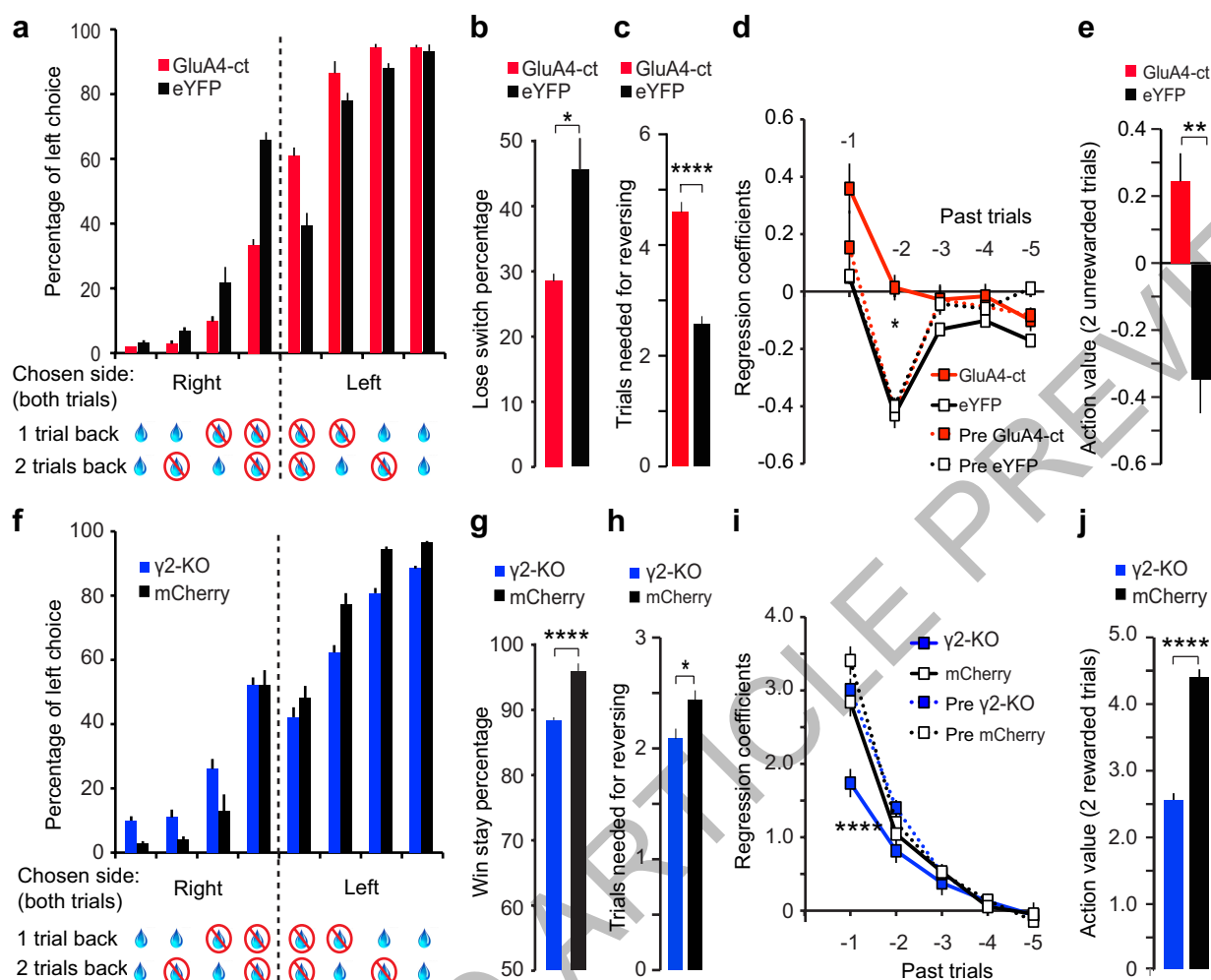
$t$  test. **e**, CS and US (reward) responses of an example putative GPh neuron tracked over multiple sessions. Session-by-session waveform correlations for this individual unit were  $>0.97$ . **f**, CS-US (reward) response index for 60 putative GPh neurons (9 mice) across different stages of training (the blue dots represent values of the sample neuron in **e**) ( $r^2 = 0.48$ ,  $P < 0.0001$  by a linear regression). **g**, Responses of an example GPh neuron to an expected reward (blue), an unexpectedly omitted reward (light blue), or an unsignalled airpuff (grey). **h**, auROC analysis of differences in firing rate between baseline and US presentation time (1.7–1.9 s) in reward omission trials ( $n = 15$  neurons from 4 mice). Filled bars,  $P < 0.05$ ,  $t$  test.





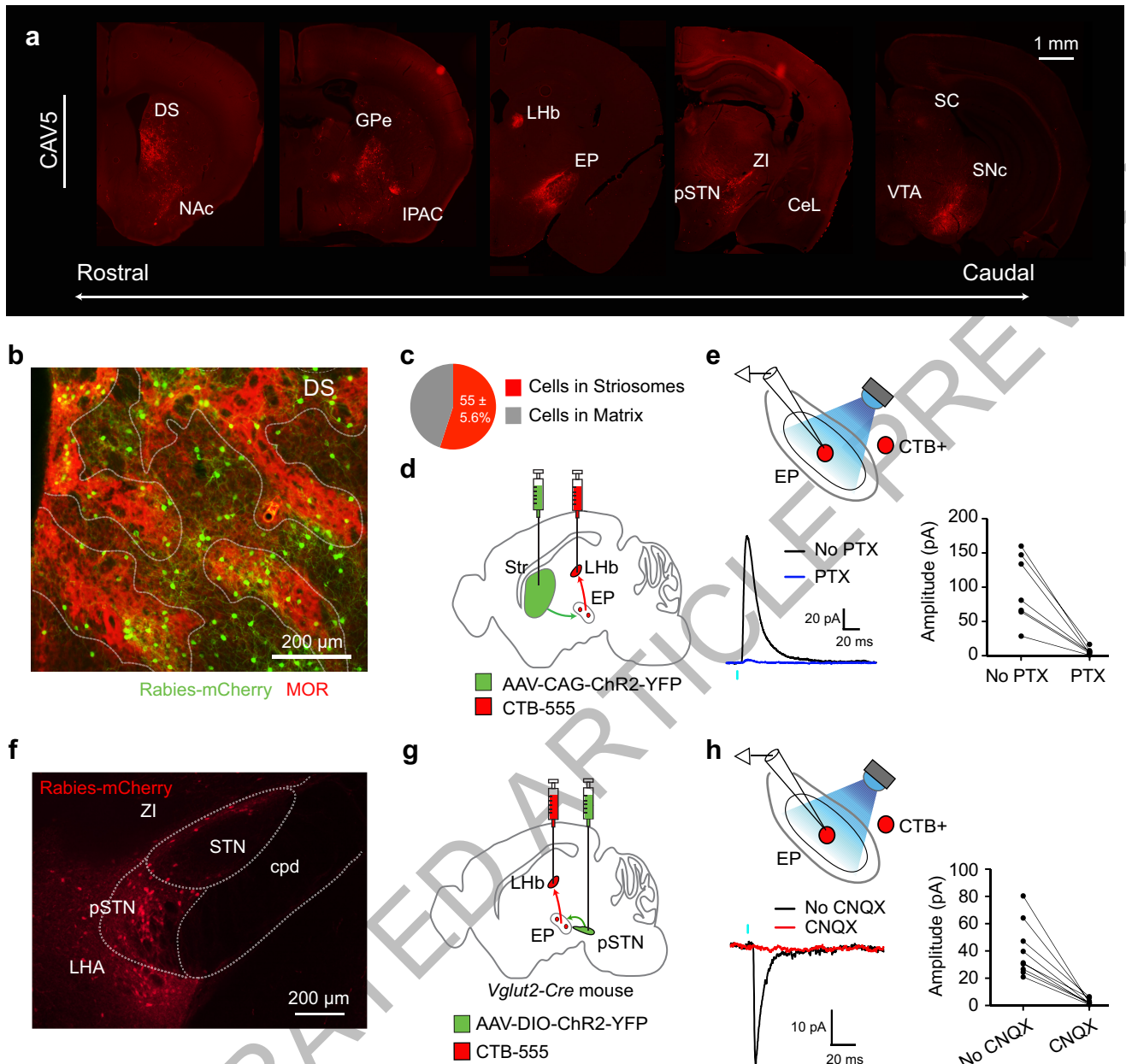
**Figure 3 | Optogenetic inhibition or activation of the GPh-LHb pathway bidirectionally influences reinforcement.** **a**, Schematic of the optogenetic inhibition. **b**, Probability of left port choice as a function of action value, for trials immediately following the trials in which photo-inhibition was delivered when mice entered the left (“stim left”) or right (“stim right”) port, or was not delivered (“no stim”). Coloured lines indicate the fit by the logistic regression model on the pooled data for each of the three conditions; grey lines indicate the pooled data for each individual mouse

( $n = 5$  mice, 34,627 trials,  $6,943 \pm 1330$  trials per mouse). **c**, Similar to **b**, except that control mice with eYFP-expressing GPh neurons were used ( $n = 6$  mice, 79,589 trials,  $13,265 \pm 596$  trials per mouse). **d** & **e**, Similar to **a** and **b**, except that optogenetic activation of the GPh-LHb projection was applied ( $n = 6$  mice, 42,292 trials,  $4,424 \pm 1806$  trials per mouse). **f**, Similar to **e**, except that control mice with eYFP-expressing GPh neurons were used ( $n = 6$  mice, 45,389 trials,  $7,564 \pm 2120$  trials per mouse). In **b**, **c**, **e**, and **f**,  $P$  values are reported for  $t$  tests:  $H_0: \beta_{\text{stim}} = 0$ .



**Figure 4 | Reducing glutamatergic or GABAergic drive onto GPh neurons decreases sensitivity to negative or positive feedback, respectively.** **a**, Bar graphs showing the increased perseverance of  $GPh^{GluA4-ct}$  mice ( $n = 10$ ) compared to  $GPh^{eYFP}$  controls (in which eYFP was introduced into GPh neurons by a Cre-dependent virus;  $n = 7$ ). For clarity, only choices for which mice had previously made two consecutive responses at the same port are shown. **b**, The lose-switch percentage in these mice ( $GPh^{GluA4-ct}$ ,  $31.13 \pm 1.7\%$ ;  $GPh^{eYFP}$ ,  $45.67 \pm 5.1\%$ ;  $T_{(13)} = 2.58$ ,  $*P < 0.05$ ,  $t$  test). **c**, The number of trials mice took before reversing choice after reward contingencies were switched ( $GPh^{GluA4-ct}$ ,  $4.59 \pm 0.18$  trials;  $GPh^{eYFP}$ ,  $2.56 \pm 0.15$  trials;  $T_{(13)} = 6.02$ ,  $****P < 0.0001$ ,  $t$  test). **d**, The negative regression coefficients associated with the past five trials for  $GPh^{GluA4-ct}$  mice and  $GPh^{eYFP}$  mice before and after surgery (first two trials back x groups,  $F_{(3,33)} = 6.566$ ,  $P < 0.01$ ;  $*P < 0.05$  for  $GPh^{GluA4-ct}$  compared to all other groups on the second trial back; two-way ANOVA followed by Bonferroni's test). **e**, The action value of two sequentially unrewarded trials, derived from the sum of their regression coefficients ( $T_{(16)} = 3.46$ ,

$**P < 0.01$ ,  $t$  test). **f**, Bar graphs showing decreased perseverance in  $GPh^{\gamma2-KO}$  mice ( $n = 8$ ) compared to  $GPh^{mCherry}$  controls (in which mCherry was introduced into GPh neurons by a Flp-dependent virus;  $n = 8$ ). Only choices where mice previously made two consecutive responses at the same port are shown. **g**, The win-stay percentage in these mice ( $GPh^{\gamma2-KO}$ ,  $89.0 \pm 0.7\%$ ;  $GPh^{mCherry}$ ,  $95.6 \pm 0.5\%$ ;  $T_{(16)} = -6.61$ ,  $****P < 0.0001$ ,  $t$  test). **h**, The number of trials mice took before reversing choice after reward contingencies were switched ( $GPh^{\gamma2-KO}$ ,  $2.08 \pm 0.26$  trials;  $GPh^{mCherry}$ ,  $2.45 \pm 0.27$  trials;  $T_{(16)} = -2.74$ ,  $*P < 0.05$ ,  $t$  test). **i**, The positive regression coefficients associated with the past five trials for  $GPh^{\gamma2-KO}$  mice and  $GPh^{mCherry}$  mice before and after surgery (first two trials back x groups,  $F_{(3,31)} = 42.10$ ,  $P < 0.0001$ ;  $****P < 0.0001$  for  $GPh^{\gamma2-KO}$  compared to all other groups on the first trial back; two-way ANOVA followed by Bonferroni's test). **j**, The action value of two sequentially rewarded trials, derived from the sum of their regression coefficients ( $T_{(16)} = -7.49$ ,  $****P < 0.0001$ ,  $t$  test). All data are represented as mean  $\pm$  s.e.m.



**Figure 5 | Identification of monosynaptic inputs to the GPe.** **a**, Series of coronal sections, ipsilateral to site of injection, from a representative mouse (i.e., CAV5) showing the major monosynaptic inputs to the GPe. **b**, Confocal image of the dorsal striatum (DS) with monosynaptically labelled neurons (green) and Mu Opioid Receptor (MOR) immunostaining that labels the striosomes (red). **c**, Quantification of monosynaptically labelled cells in striatal subcompartments. **d**, The injection strategy. **e**, Schematic of recording configuration (upper) and sample inhibitory postsynaptic currents (IPSCs) induced by optogenetic

activation of the striatal input to the GPe (lower left). These IPSCs were blocked by picrotoxin (PTX) (lower right). **f**, Image of the pSTN with monosynaptically labelled neurons (red). **g**, The injection strategy. **h**, Schematic of recording configuration (upper) and sample excitatory postsynaptic currents (EPSCs) induced by optogenetic activation of pSTN input to the GPe (lower left). These EPSCs were blocked by CNQX (lower right). Diagrams in **d** and **g** are modified from the Allen Mouse Brain Atlas, Allen Institute for Brain Science; available from <http://mouse.brain-map.org/>.



## METHODS

**Animals.** All procedures were approved by the Institutional Animal Care and Use Committee of Cold Spring Harbor Laboratory (CSHL) and conducted in accordance to the United States' National Institutes of Health guidelines. Mice were housed under a 12h light-dark cycle (8 a.m. to 8 p.m. light). All behavioural experiments were performed during the light cycle. All mice had free access to food, but water was restricted for mice used in certain behavioural experiments. Free water was provided on days with no experimental sessions. Male and female mice 2–4 months of age were used in all experiments. No differences were observed in the behavior of male or female mice during the switching task or in our optogenetic or synaptic manipulations for this behavior (see below). All animals were randomly allocated to the different experimental conditions used in this study. The *Vglut2-Cre* (*Slc17a6<sup>tm2(Cre)Low</sup>/J*, stock #016963 from Jackson Laboratory, Bar Harbor, Maine, USA), *Ai35* (*Gt(ROSA)26Sor<sup>tm35.1(CAG-aop3/GFP)Hze</sup>/J*, stock #012735 from Jackson Laboratory), *Rosa26-stop<sup>fllox</sup>-H2b-GFP* (from Dr. Z. Josh Huang, CSHL)<sup>31</sup>, *Som-Flp* (from Dr. Z. Josh Huang)<sup>32</sup>, *Gabrg2<sup>fllox</sup>* (*Gabrg2<sup>tm1Wid</sup>/J*, stock #021197 from Jackson Laboratory)<sup>23</sup>, *Rosa26-stop<sup>fllox</sup>-tTA* (stock #012266 from Jackson Laboratory)<sup>32,33</sup> mouse strains have all been previously characterized. All mice were bred onto a C57BL/6J background.

**Viral vectors.** All adeno-associated viruses (AAV) were produced by the University of North Carolina vector core facility (Chapel Hill, North Carolina, USA) or the University of Pennsylvania vector core (Pennsylvania, USA) and have previously been described: AAV9-Efla-DIO-hChR2(H134R)-eYFP, AAV9-CAG-FLEX-ArchT-GFP, AAV9-Efla-DIO-eYFP, AAV8-hSyn-DIO-mCherry, AAV1-Syn-GCAMP6f.WPRE.SV40 (used for non-Cre dependant viral tracing), AAV9-CAG-ChR2-GFP, AAV9-DIO-GluA4-ct-GFP<sup>22</sup>, AAV9-CAG-FSF-GFP-T2A-nCre (which expresses Cre in a Flp-dependent manner<sup>22</sup>), and AAV8-Efla-DIO-2A-mCherry (which expresses mCherry in a Flp-dependent manner<sup>22</sup>). CAV2-Cre was purchased from Montpellier vector platform (Plateforme de Vectorologie de Montpellier (PVM), Biocampus Montpellier, Montpellier, France). All viral vectors were aliquoted and stored at  $-80^{\circ}\text{C}$  until use.

**Stereotaxic surgery.** Mice were anesthetized with  $100\text{ mg kg}^{-1}$  /  $0.4\text{ mg kg}^{-1}$  ketamine / dexmedetomidine hydrochloride and head-fixed in a stereotaxic injection frame (myNeuroLab, Leica Microsystems Inc., Buffalo Grove, Illinois, USA). Lidocaine ( $20\mu\text{l}$ ) was injected subcutaneously into the head and neck area as a local anaesthetic. For *in vivo* recordings, mice were implanted with a head-bar and a microdrive containing the recording electrodes and an optical fibre. Viral injections were performed using previously described procedures<sup>32</sup> at the following stereotaxic coordinates: ventral medial nucleus of the thalamus (VM),  $-1.4$ – $-1.5\text{ mm}$  from bregma,  $1.3\text{ mm}$  lateral from midline, and  $4.10\text{ mm}$  ventral from cortical surface; GPh,  $-1.22\text{ mm}$  from bregma,  $1.77\text{ mm}$  lateral from midline, and  $4.64\text{ mm}$  ventral from cortical surface; and LHB,  $-1.7\text{ mm}$  from bregma,  $0.53\text{ mm}$  lateral from midline, and  $2.8\text{ mm}$  ventral from cortical surface. During the surgical procedure, mice were kept on a heating pad and were brought back to their home-cage for post-surgery recovery and monitoring. Postoperative care included intraperitoneal injection with  $0.3$ – $0.5\text{ ml}$  of Lactated Ringer's solution and Metacam ( $1$ – $2\text{ mg kg}^{-1}$  meloxicam; Boehringer Ingelheim Vetmedica, Inc., St. Joseph, Missouri, USA) for analgesia and anti-inflammatory purposes. All AAVs were injected at a total volume of approximately  $0.6\mu\text{l}$ , and were allowed at least 4 weeks for maximal expression. For retrograde tracing of projection cells in the EP, CTB-555 or CTB-488 ( $0.3\mu\text{l}$ ,  $0.5\%$  in phosphate-buffered saline (PBS); Invitrogen, Thermo Fisher Scientific, Waltham, Massachusetts, USA) was injected into the VM or the LHB and allowed 3–5 days for sufficient retrograde transport.

**Immunohistochemistry.** Immunohistochemistry experiments were performed following standard procedures. Briefly, mice were anesthetized with Euthasol ( $0.4\text{ ml}$ ; Virbac, Fort Worth, Texas, USA) and transcardially perfused with  $40\text{ ml}$  of PBS, followed by  $40\text{ ml}$  of  $4\%$  paraformaldehyde in PBS. Coronal sections ( $40$ – $50\mu\text{m}$ ) were cut using a freezing microtome (Leica SM 2010R, Leica). Sections were first washed in PBS ( $3 \times 5\text{ min}$ ), incubated in PBST ( $0.3\%$  Triton X-100 in PBS) for  $30\text{ min}$  at room temperature (RT) and then washed with PBS ( $3 \times 5\text{ min}$ ). Next, sections were blocked in  $5\%$  normal goat serum in PBST for  $30\text{ min}$  at RT and then incubated with primary antibodies overnight at  $4^{\circ}\text{C}$ . Sections were washed with PBS ( $5 \times 15\text{ min}$ ) and incubated with fluorescent secondary antibodies at RT for  $1\text{ h}$ . After washing with PBS ( $5 \times 15\text{ min}$ ), sections were mounted onto slides with Fluoromount-G (eBioscience, San Diego, California, USA). Images were taken using a LSM 710 laser-scanning confocal microscope (Carl Zeiss, Oberkochen, Germany). The primary antibodies used were: rabbit anti-Somatostatin-14 (Peninsula Laboratories Inc., San Carlos, California, USA; catalogue number T-4103), mouse anti-Parvalbumin (Swant, Switzerland; PV 235), chicken anti-GFP (Aves Labs, catalogue number GFP1020, lot number GFP697986), rabbit anti-RFP (Rockland, catalogue number 600-401-379, lot number 34135). Primary antibodies were incubated with appropriate

fluorophore-conjugated secondary antibodies (Life Technologies, Carlsbad, California, USA) depending on the desired fluorescence colour.

**Monosynaptic tracing with pseudotyped rabies virus.** Retrograde tracing of mono-synaptic inputs onto genetically-defined cell populations of the GPh was accomplished using a previously described method<sup>5,32</sup>. In brief, *Vglut2-Cre; Rosa26-stop<sup>fllox</sup>-tTA* mice that express tTA in *Vglut2<sup>+</sup>* cells were injected into the GPh with AAV-TRE-hGFP-TVA-G ( $0.2$ – $0.3\mu\text{l}$ ) that expresses the following components in a tTA-dependent manner: a fluorescent reporter histone GFP (hGFP); TVA (which is a receptor for the avian virus envelope protein EnvA); and the rabies envelope glycoprotein (G). Alternatively, CAV2-Cre virus was injected into the LHB of *Rosa26-stop<sup>fllox</sup>-tTA* mice, so that any input to the LHB will express tTA. As above these mice were also injected into the GPh with AAV-TRE-hGFP-TVA-G ( $0.2$ – $0.3\mu\text{l}$ ). Two weeks later, mice were injected in the same GPh location with the rabies-EnvA-SAD-DG-mCherry ( $1.2\mu\text{l}$ ), a rabies virus that is pseudotyped with EnvA, lacks the envelope glycoprotein, and expresses mCherry. This method ensures that the rabies virus exclusively infects cells expressing TVA. Furthermore, complementation of the modified rabies virus with envelope glycoprotein in the TVA-expressing cells allows the generation of infectious particles, which then can trans-synaptically infect presynaptic neurons.

**Center of mass analysis.** To compare average location of rabies infected neurons in the GPe and STN, the center of mass of a brain section was obtained by averaging positions of neurons. In order to standardize the results from individual animals onto a standard atlas, each neuron's position was normalized by anatomical landmarks: for Extended Data Figure 9g, we used the midline and the most ventral part of the GPe; for Extended Data Figure 9i, we used the midline and the most ventral medial portion of the STN.

**Classical conditioning task.** Nine *Vglut2-Cre; Ai35* mice were trained on an auditory classical conditioning task. One week after surgery mice were water-deprived in their home-cage. During training, mice were head restrained using custom-made clamps and metal head-bars. Each mouse was habituated to head restraint for one day prior to training. There were five possible outcomes (unconditioned stimuli, US), each associated with a different auditory cue (conditioned stimulus, CS): a large water reward ( $5\mu\text{l}$ ), a small water reward ( $1\mu\text{l}$ ), nothing, a small air puff ( $100\text{ ms}$ ) or a large air puff ( $500\text{ ms}$ ). The air puff was delivered to the animal's face. Each trial began with a CS ( $1\text{ second}$  sound), followed by a  $0.5\text{ second}$  delay and then a US (the outcome). In each session, reward and punishment trials were presented in two sequential blocks, with each cue chosen pseudorandomly. Each block contained the neutral stimulus.

Eye blinking was tracked using a CMOS camera (QSIIC2). Offline video analysis was conducted using EthoVision XT software (Noldus; Wageningen, The Netherlands). Oval regions of interest (ROI) surrounding the eye were manually drawn. Pixels corresponding to the eye were detected as those that were darker than the background within the ROI. As each blink reduced the observable area of the eye, a threshold number of pixels corresponding to the eye was used to define a blink, and thus to determine the time and duration of each blink.

**In vivo electrophysiology.** Custom-built screw-driven microdrives with 4 implantable tetrodes and a  $50\mu\text{m}$  fibre-optic were used to record simultaneously from multiple neurons. Each tetrode was glued to the fibre-optic with epoxy, such that the end of each tetrode was  $200$ – $400\mu\text{m}$  from the end of the fibre-optic. Neural recordings and time stamps for behavioural variables were acquired with a Tucker-Davis Technologies RZ recording system (with a 32 channel preamp PZ2-32 and a RZ5D Bioamp processor; Alachua, Florida, USA).

Broadband signals from each wire were filtered between  $0.2$  and  $8,500\text{ Hz}$  and recorded continuously at  $25\text{ kHz}$ . To extract the timing of spikes, signals were band-pass-filtered between  $300$ – $5,000\text{ Hz}$ . Data analyses were carried out using software in Matlab (The Mathworks, Inc., Natick, Massachusetts, USA). Spike waveforms were manually sorted offline based on amplitude and waveform energy features using MClust-3.5 (from Dr. A. David Redish, University of Minnesota, Minneapolis, Minnesota, USA). Individual neurons were only included in the dataset if they were well isolated based on their isolation distance ( $>20$ ) and L-ratio ( $<0.1$ )<sup>34</sup>. Prior to implantation, tetrodes were dipped in DiI to aid the post-hoc visualization of the recording locations.

In order to convert raster plots of firing rate into continuous spike density functions, spike times were first binned into  $1\text{ ms}$  time windows and then convolved with a Gaussian kernel ( $\sigma = 15\text{ ms}$ ). To determine the response to the CS or US presentation, the average firing rates were calculated using a  $300\text{ ms}$  window defined as  $180$ – $480\text{ ms}$  following the stimulus. These time windows were chosen to cover the time of the peak neuronal response. Average baseline firing was calculated using a  $300\text{ ms}$  window immediately preceding the delivery of the CS.

To identify putative GPh neurons – the *Vglut2<sup>+</sup>* EP neurons – we used Arch-mediated optic tagging<sup>35</sup>, whereby  $200\text{ ms}$  light pulses ( $\lambda = 532\text{ nm}$ ; OEM Laser Systems Inc., Bluffdale, Utah, USA) were delivered every  $3\text{ seconds}$  for  $100\text{ trials}$

following each behavioural recording session. In early sessions we also used 500 ms ( $n = 3$ ) or 1 second ( $n = 1$ ) light pulses, which tagged Vglut2<sup>+</sup> EP neurons in a similar way to that of the 200 ms light pulses.

In addition to their response to light, putative GPh neurons were identified based on their firing pattern through a previously described unsupervised clustering approach<sup>15</sup>. Briefly, to calculate receiver-operating characteristic (ROC) curves, the distribution of firing rates within 100 ms bins were compared (from 1 second prior to the CS presentation to 1.5 seconds after delivery of the US) to the baseline firing rate 900 ms prior to CS presentation. The first three principal components (PCs) of the auROC (area under the receiver operating characteristic) curves were then calculated using principal component analysis (PCA), with the singular value decomposition algorithm. Hierarchical clustering of the auROC curves' first three PCs was then performed using a Euclidean distance metric and a complete agglomeration method.

Cross-correlations between spike waveforms across sessions were used to determine whether the same unit was recorded over multiple sessions. The cross-correlations were calculated after aligning the negative peak of each waveform, averaging separately, and aligning the peaks of the averages. A conservative session-to-session cross-correlation coefficient of  $>0.95$  was used to positively classify two sets of waveforms as belonging to the same unit. The correlation was calculated using the full duration of the spike in a window 10 ms prior to and 40 ms after the peak negative response.

CS-US indices were calculated as  $(CS - US)/(CS + US)$ , where CS is the difference between the peak firing rate (maximum value of the PSTH) in the 500 ms after CS onset and the baseline firing rate, and US is the difference between the peak firing rate in the 500 ms after US onset and the baseline firing rate. The baseline firing rate was calculated as the mean of the PSTH in the 0.5 s before CS onset.

**Probabilistic switching task.** Mice were trained in a two alternative choice probabilistic switching task. The initiation port was located in the centre between two reward ports. Infrared photodiode/phototransistor pairs placed on the inside of each port detected each nose poke (Island Motion Corporation, Tappan, New York, USA). Water valves (Neptune Research & Development, Inc., West Caldwell, New Jersey, USA) were calibrated to deliver a volume of water (2  $\mu$ l) for rewarded choices. Water-deprived mice initiated a trial by a nose poke into the central port, which triggered a 'Go' light cue over the two peripheral ports. Mice then chose to enter either the left or right peripheral port where they received water rewards. On each trial, reward was delivered only at one port, and only for 75% of correct choices. The rewarded port was switched across blocks – the lengths of which were randomly distributed between 7-23 rewards – with no external instruction.

To characterize how reward and lack of reward influenced choice on a trial-by-trial basis in this task, we used a previously established logistic regression model<sup>21,36</sup>

$$\log\left(\frac{P_L(i)}{1 - P_L(i)}\right) = \sum_{j=1}^n \beta_j^{\text{Reward}} (Y_L(i-j) - Y_R(i-j)) + \sum_{j=1}^n \beta_j^{\text{No Reward}} (N_L(i-j) - N_R(i-j)) + \beta_0 \quad (1)$$

where  $P_L(i)$  indicates the probability of choosing the left port on the  $i$ -th trial;  $Y_L(i)$  and  $Y_R(i)$  indicate a reward was delivered when choosing the left or right port, respectively, on the  $i$ -th trial (1 for chosen and 0 for non-chosen port);  $N_L(i)$  and  $N_R(i)$  specify the lack of reward when choosing the left or right port, respectively, on the  $i$ -th trial (1 for chosen and 0 for non-chosen port);  $n$  represents the number of past trials that were included in the regression model ( $n = 5$  was used, except for the optogenetic experiments where  $n = 2$  was used); the regression coefficients  $\beta^{\text{Reward}}$  and  $\beta^{\text{No Reward}}$  represent the weighted contributions past rewards or lack of rewards have to the current choice;  $\beta_0$  indicates the intrinsic bias a mouse may have for choosing the left or right port.

For *in vivo* optogenetic manipulations in the probabilistic switching task, Vglut2-cre mice were bilaterally implanted with optical fibre cannulae (Thorlabs, Inc., Newton, New Jersey, USA), prior to behavioural training and following the surgery procedure for viral injection (described above). Optical fibres (200  $\mu$ m) were implanted with the tips placed 0.4 mm dorsal to the site of virus injection and were secured to the skull with C&B Metabond quick adhesive cement (Parkell Inc., Edgewood, New York, USA) followed by dental cement (Lang Dental Manufacturing Co., Inc., Wheeling, Illinois, USA). Viruses were allowed to express for 3-4 weeks. The optic fibres were connected to a laser source ( $\lambda = 532$  nm or 473 nm; OEM Laser Systems) via a dual fibre rotary joint (FRJ\_1x2i\_FC-2FC; Doric Lenses, Inc., Québec, Canada) using an optic fibre sleeve (Thorlabs). Following training and habituation, optical stimulation was delivered at two time points during the task, at the time of action selection when the mouse nose pokes in the centre port to initiate a trial or during the evaluation phase when the mouse nose pokes in the peripheral choice ports. For ChR2-mediated stimulation, 5 ms

optical light pulses were delivered at 30 Hz for 500 ms. For Arch-mediated inhibition, 500 ms of continuous illumination was delivered. In each session stimulation occurred randomly at either choice port in 10% of the trials. Stimulation sessions were interspersed by training sessions.

The effect of optogenetic manipulation on outcome evaluation was model by the following logistic regression equation

$$\log\left(\frac{P_L(i)}{1 - P_L(i)}\right) = \sum_{j=1}^n \beta_j^{\text{Reward}} (Y_L(i-j) - Y_R(i-j)) + \sum_{j=1}^n \beta_j^{\text{No Reward}} (N_L(i-j) - N_R(i-j)) + \beta_0 + \beta_{\text{stim}} \times X_{\text{stim}}(i-1) \quad (2)$$

where  $\beta_{\text{stim}}$  is added to equation 1 to represent the effects of photo-stimulation on the current choice, and  $X_{\text{stim}}$  represents whether in the previous trial the stimulation was delivered when mice nose poked the left (1) or right (-1) reward port to collect reward, or was not delivered (0). When analysing the longevity of the effect optogenetic stimulation had on upcoming choices, additional  $\beta_{\text{stim}}$  terms were added to equation 2 to account for the stimulation on the  $(i-n)$  trials.

The effect of optogenetic manipulations on action selection was model by the following logistic regression equation

$$\log\left(\frac{P_L(i)}{1 - P_L(i)}\right) = \sum_{j=1}^n \beta_j^{\text{Reward}} (Y_L(i-j) - Y_R(i-j)) + \sum_{j=1}^n \beta_j^{\text{No Reward}} (N_L(i-j) - N_R(i-j)) + \beta_0 + \beta_{\text{stim}} \times X_{\text{stim}}(i) \quad (3)$$

where  $\beta_{\text{stim}}$  is added to equation 1 to represent the effects of photo-stimulation on the current choice, and  $X_{\text{stim}}$  represents whether the stimulation was delivered in the current trial when mice nose poked the centre port to initiate the trial (1 for stimulated trials and 0 otherwise).

For cell-type specific synaptic manipulations and their controls, mice were trained on the task prior to surgery and then tested again 4 weeks after surgery. The first five sessions prior to surgery and after surgery were used as the comparative sessions.

**In vitro electrophysiology.** Patch clamp recording was performed as previously described<sup>32</sup>. Briefly, mice were anesthetized with isoflurane before they were decapitated; their brains were then dissected out and placed in ice chilled dissection buffer (110 mM choline chloride, 25 mM NaHCO<sub>3</sub>, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub>, 2.5 mM KCl, 0.5 mM CaCl<sub>2</sub>, 7.0 mM MgCl<sub>2</sub>, 25.0 mM glucose, 11.6 mM ascorbic acid and 3.1 mM pyruvic acid, gassed with 95% O<sub>2</sub> and 5% CO<sub>2</sub>). An HM650 Vibrating-blade Microtome (Thermo Fisher Scientific) was then used to cut 300  $\mu$ m thick coronal sections that contained the EP. These slices were subsequently transferred to a storage chamber that contained oxygenated artificial cerebrospinal fluid (ACSF) (118 mM NaCl, 2.5 mM KCl, 26.2 mM NaHCO<sub>3</sub>, 1 mM NaH<sub>2</sub>PO<sub>4</sub>, 20 mM glucose, 2 mM MgCl<sub>2</sub> and 2 mM CaCl<sub>2</sub>, at 34 °C, pH 7.4, gassed with 95% O<sub>2</sub> and 5% CO<sub>2</sub>). Following 40 min of recovery time, slices were transferred to RT (20-24 °C), where they were continuously bathed in the ACSF.

Visually guided whole-cell patch clamp recording from GPh neurons was obtained with Multiclamp 700B amplifiers and pCLAMP 10 software (Molecular Devices, Sunnyvale, California, USA), and was guided using an Olympus BX51 microscope equipped with both transmitted and epifluorescence light sources (Olympus Corporation, Shinjuku, Tokyo, Japan). GPh neurons with fluorescence of different colours were patched. To evoke excitatory postsynaptic currents (EPSCs), a bipolar stimulating electrode was placed on the medial dorsal border of the EP. Electrical stimulation was delivered every 10 seconds and synaptic responses were low-pass filtered at 1 KHz and recorded at holding potentials of -70 mV (for AMPA-receptor-mediated responses) and +40 mV (for NMDA-receptor-mediated responses). The NMDA-receptor-mediated component of the response was quantified as the mean current amplitude between 50-60 ms after electrical stimulation. Recordings were made in ACSF. The internal solution contained 115 mM caesium methanesulphonate, 20 mM CsCl, 10 mM HEPES, 2.5 mM MgCl<sub>2</sub>, 4 mM Na<sub>2</sub>ATP, 0.4 mM Na<sub>3</sub>GTP, 10 mM sodium phosphocreatine and 0.6 mM EGTA (pH 7.2). The evoked EPSCs were recorded with picrotoxin (100  $\mu$ M) added to the ACSF,

and were analyzed using pCLAMP 10 software. Miniature inhibitory postsynaptic currents (mIPSCs) were recorded at 0 mV holding potential with tetrodotoxin (TTX; 1  $\mu$ M), APV (100  $\mu$ M), and CNQX (5  $\mu$ M) added to the ACSF, and were analysed using Mini Analysis software (Synaptosoft, Inc., Decatur, Georgia, USA).

To evoke striatal or pSTN synaptic transmission onto GPh neurons, AAV-ChR2-YFP or AAV-DIO-ChR2-YFP was injected into the striatum of wild-type C57BL/6 mice or the pSTN of *Vglut2-Cre* mice, respectively, and allowed to express for 3 weeks. Acute brain slices were prepared and a blue light was used to stimulate ChR2-expressing axons. The light source was a single-wavelength LED system ( $\lambda = 470$  nm; <http://www.cooled.com/>) connected to the epifluorescence port of the Olympus BX51 microscope. Single light pulses of 1 ms, triggered by a TTL signal from the Clampex software, were delivered to drive synaptic responses.

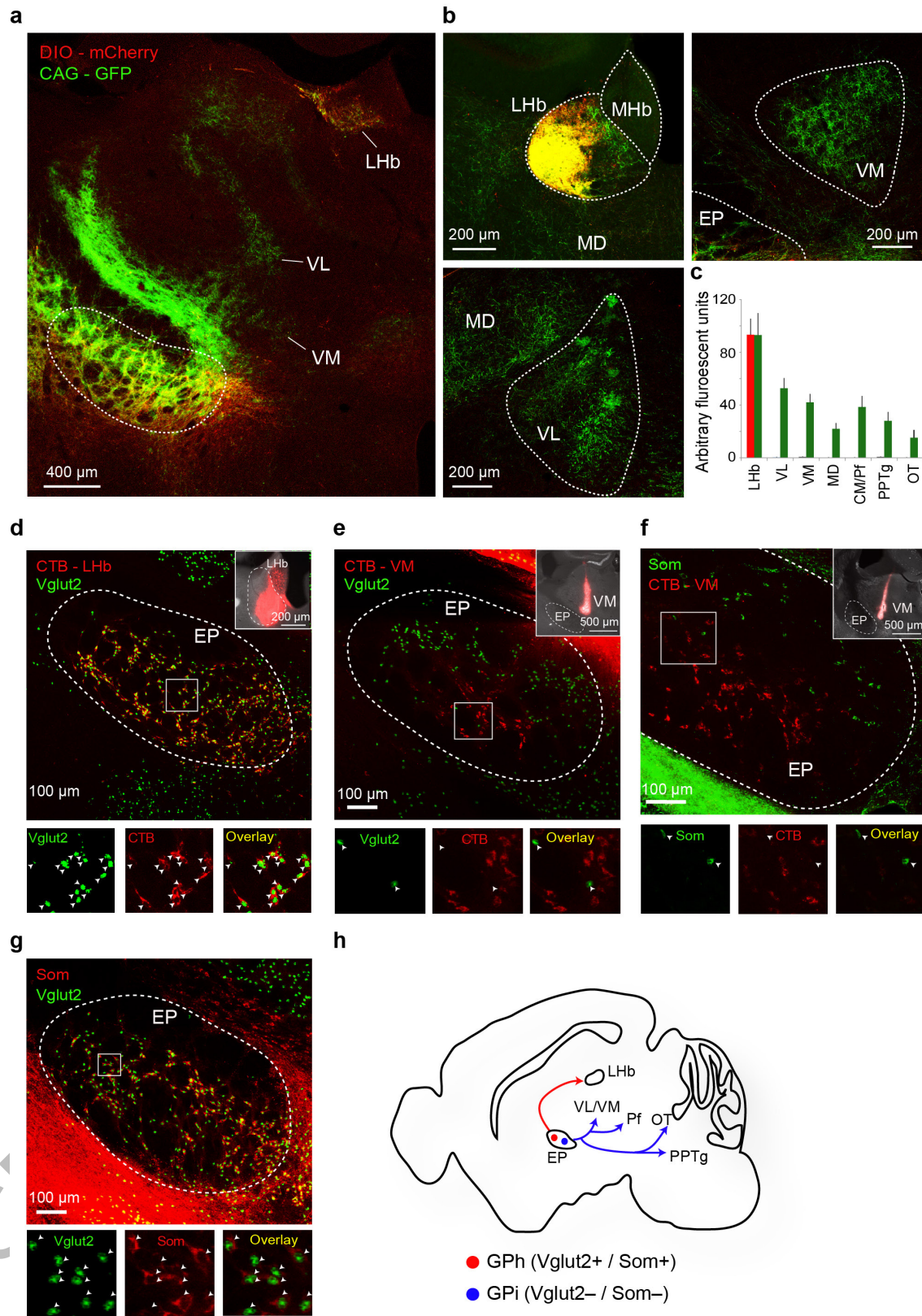
**Statistics and data presentation.** To determine whether parametric tests could be used, the Shapiro-Wilk Test was performed on all data as a test for normality. The statistical test used for each comparison is indicated when used. The sample sizes used in this study were based on estimations by a power analysis. Behavioural tests and electrophysiological data acquisition were performed by investigators with knowledge of the identities of experimental groups. All these experiments were

controlled by computer systems, with data collected and analysed in an automated and unbiased way. For *in vivo* recordings, the data from a mouse were excluded if the tetrode tracts and tips were outside of the EP. No other mice or data points were excluded.

31. He, M. *et al.* Cell-type-based analysis of microRNA profiles in the mouse brain. *Neuron* **73**, 35–48, doi:10.1016/j.neuron.2011.11.010 (2012).
32. Penzo, M. A. *et al.* The paraventricular thalamus controls a central amygdala fear circuit. *Nature* **519**, 455–459, doi:10.1038/nature13978 (2015).
33. Li, L. *et al.* Visualizing the distribution of synapses from individual neurons in the mouse brain. *PloS one* **5**, e11503, doi:10.1371/journal.pone.0011503 (2010).
34. Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K. & Redish, A. D. Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* **131**, 1–11, doi:10.1016/j.neuroscience.2004.09.066 (2005).
35. Courtin, J. *et al.* Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. *Nature* **505**, 92–96, doi:10.1038/nature12755 (2014).
36. Lau, B. & Glimcher, P. W. Dynamic response-by-response models of matching behavior in rhesus monkeys. *Journal of the experimental analysis of behavior* **84**, 555–579 (2005).

ACCELERATED ARTICLE PREVIEW

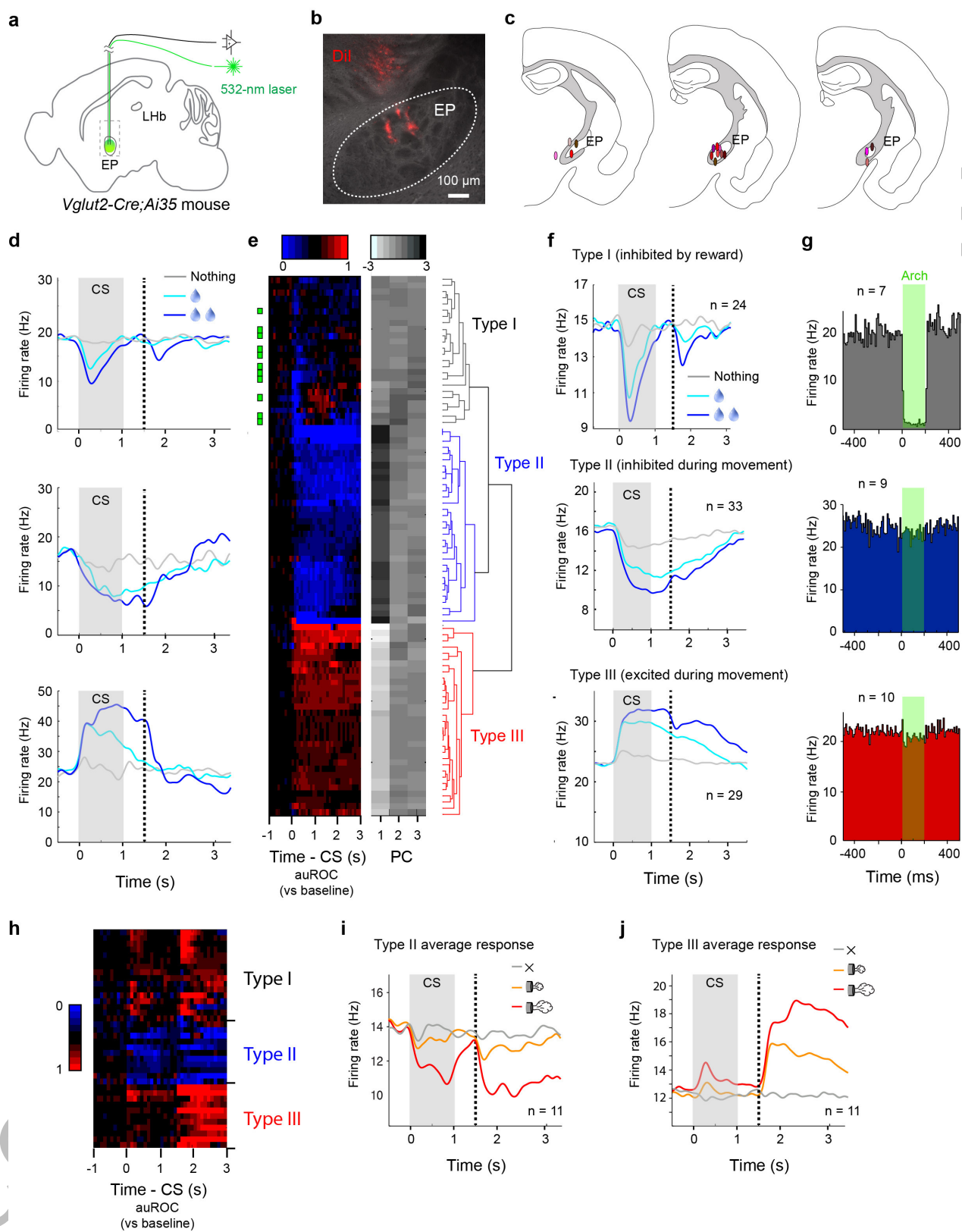




Extended Data Figure 1 | See next page for caption.

**Extended Data Figure 1 | Vglut2 and Somatostatin are markers for GPh neurons.** **a**, Image showing the projection patterns of nonspecifically labelled neurons (green, infected with adeno-associated virus (AAV) expressing GCaMP6 (AAV1-Syn-GCaMP6f.WPRE.SV40); signal was enhanced by anti-GFP antibody; see Methods) and Vglut2<sup>+</sup> neurons (red, infected with AAV expressing mCherry in a Cre-dependent manner (AAV8-hSyn-DIO-mCherry); signal was enhanced by anti-mCherry antibody; see Methods) in the EP of a *Vglut2-cre* mouse. **b**, Confocal images of the LHb, ventrolateral thalamus (VL) and ventromedial thalamus (VM), showing fibers originating from the nonspecifically labelled neurons (green) and Vglut2<sup>+</sup> neurons (red) in the EP. **c**, Quantification of the GFP and mCherry fluorescence intensity in the projection targets of the EP neurons. **d**, Upper panel: representative image showing retrograde labelling of GPh neurons by injection of the cholera toxin subunit B conjugated to Alexa Fluor 594 (CTB-594) into the LHb (inset) of *Vglut2-cre; Rosa26-stop<sup>fllox</sup>-H2b-GFP* mice, in which Vglut2<sup>+</sup> cells can be identified based on their expression of nuclear GFP. Lower panels: high magnification pictures of the boxed area in the EP in the upper panel, showing the co-labelling of GPh neurons by CTB-594 and Vglut2 (arrowheads). The vast majority of CTB-labelled neurons expressed Vglut2 ( $95.45 \pm 1.2\%$  (mean  $\pm$  s.e.m.),  $n = 6$  mice). **e**, Upper panel: a

representative image showing retrograde labelling of VM-projecting EP neurons by injection of CTB-594 into the VM (inset) of *Vglut2-cre; Rosa26-stop<sup>fllox</sup>-H2b-GFP* mice. Lower panels: high magnification pictures of the boxed area in the EP in the upper panel, showing the segregation of the EP neurons labelled by CTB-594 and those labelled by Vglut2 (arrowheads). Very few CTB-labelled neurons expressed Vglut2 ( $0.51 \pm 0.45\%$ ,  $n = 6$  mice). **f**, Upper panel: a representative image showing retrograde labelling of VM-projecting EP neurons by injection of CTB-594 into the VM (inset). Lower panels: high magnification pictures of the boxed area in the EP in the upper panel, showing the segregation of the EP neurons labelled by CTB-594 and those labelled by anti-Som antibody (arrowheads). Very few CTB-labelled cells expressed Som ( $0.88 \pm 0.72\%$ ,  $n = 5$  mice). **g**, Upper panel: a representative image showing antibody labelling of Som in the EP of *Vglut2-Cre; Rosa26-stop<sup>fllox</sup>-H2b-GFP* mice. Lower panels: high magnification pictures of the boxed area in the upper panel, showing the co-labelling of EP neurons by Som and Vglut2 (arrowheads). The vast majority of Vglut2 neurons expressed Som ( $90.87 \pm 0.79\%$ ,  $n = 6$  mice). **h**, A cartoon showing the only projection target of GPh neurons (red) and the different projection targets of classic GPi neurons (blue). Diagram in **h** was modified from the Allen Mouse Brain Atlas, Allen Institute for Brain Science; available from <http://mouse.brain-map.org/>.



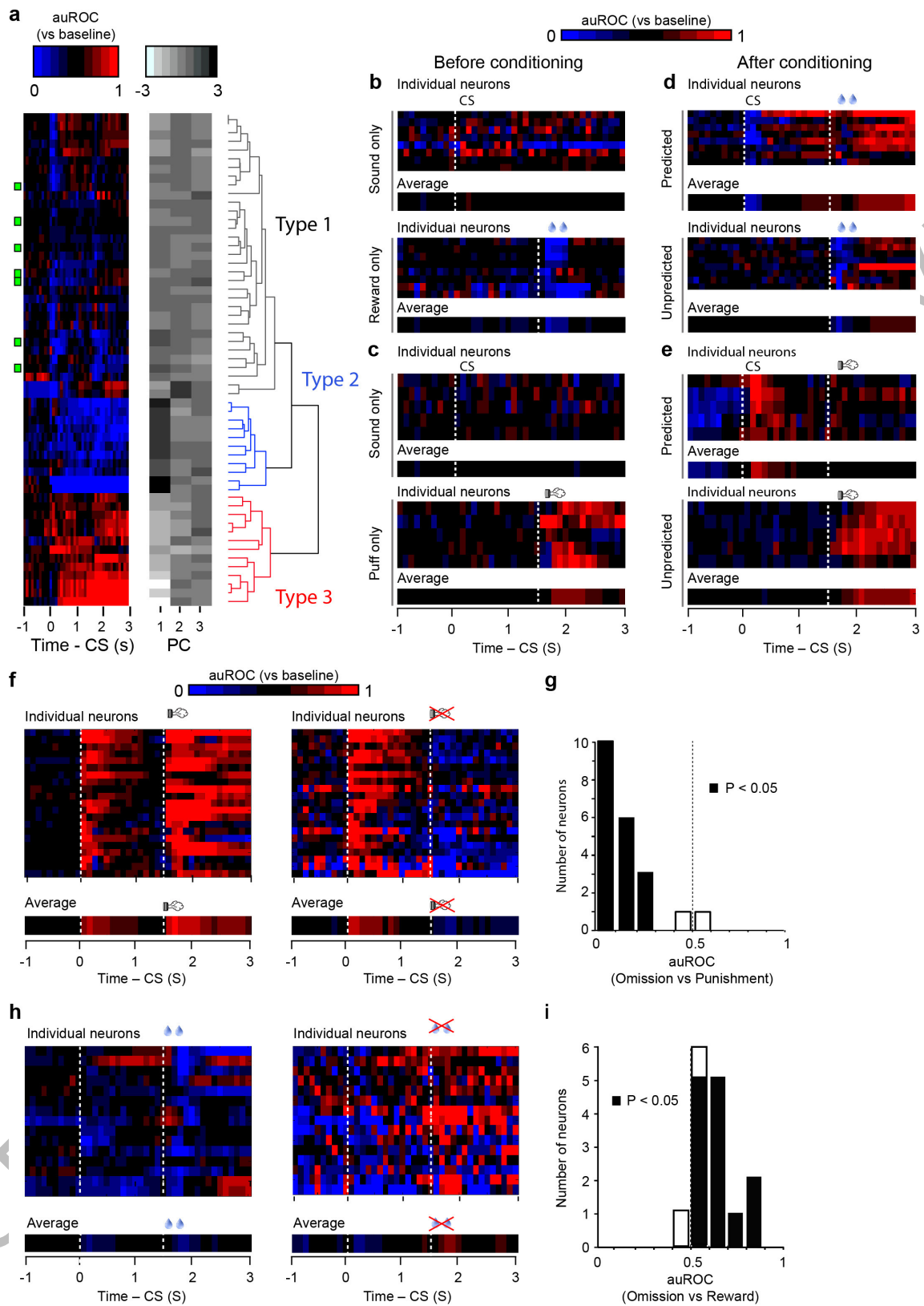
Extended Data Figure 2 | See next page for caption.



**Extended Data Figure 2 | Classification of EP neurons on the basis of their distinct response profiles.**

**a**, Schematic of the experimental approach used for *in vivo* recording and optogenetic tagging. **b**, Photomicrograph showing a DiI labelled recording site. **c**, Schematics showing the locations of the recording sites ( $n = 15$  mice). **d**, Responses of three example neurons in the classic conditioning task. **e**, Left: auROC plots of the responses of all neurons during large reward trials. Red, increase from baseline; blue, decrease from baseline; each row represents one neuron. Green bars indicate the neurons that were “optogenetically tagged” ( $n = 11$  neurons). The three main clusters are arranged in order to match the neurons presented in **d**. Right: first three principle components and hierarchical clustering dendrogram showing the relationship of each neuron within the three clusters. **f**, Average firing rates of the three types

of neurons ( $n = 86$  neurons from 9 mice). **g**, Plots of peristimulus time histogram (PSTH) showing inhibition for type I (top,  $n = 7$  neurons from 4 mice), but no change for type II (middle,  $n = 9$  neurons from 4 mice) or type III (bottom,  $n = 10$  neurons from 4 mice) neurons in response to green light pulses (green bars, 200 ms; 100 trials per neuron, 0.3 Hz). Only type II and type III neurons that were recorded in the same sessions and animals as those of the light-responsive type I neurons represented in **g** are shown. **h**, auROC plots of the responses of all 38 neurons ( $n = 9$  mice) recorded during large punishment trials. **i** & **j**, Average firing rates of type II ( $n = 11$  neurons from 9 mice) (**i**) and type III ( $n = 11$  neurons from 9 mice) (**j**) neurons during punishment trials. Diagrams in **a** and **c** were modified from the Allen Mouse Brain Atlas, Allen Institute for Brain Science; available from <http://mouse.brain-map.org/>.



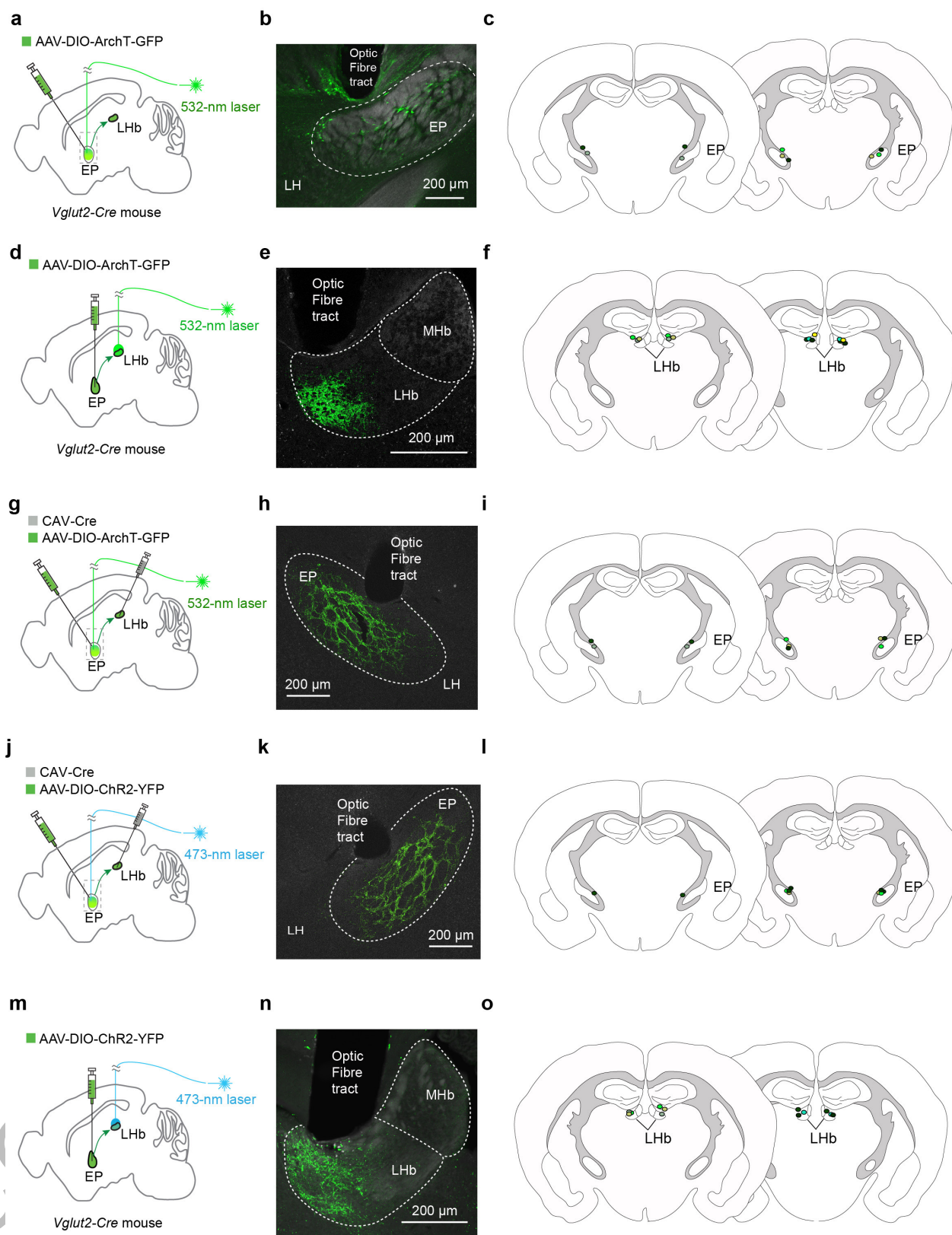
Extended Data Figure 3 | See next page for caption.

**Extended Data Figure 3 | Response profiles of putative GPh neurons during different CS-US contingencies.** **a**, Graphs showing hierarchical clustering used to identify additional putative GPh neurons used in the analysis for this figure. All data shown (b-i) are from type I neurons only. Left: auROC plots of the responses of all additional neurons recorded. Red, increase from baseline; blue, decrease from baseline. Each row represents one neuron. Green bars indicate the neurons that were optogenetically tagged. Right: first three principle components and hierarchical clustering dendrogram showing the relationship of each neuron within the three clusters. **b**, auROC plots showing the firing rate changes in response to CS (top) and reward (bottom) prior to behavioural training. **c**, auROC plots showing the firing rate changes in response to CS (top) and airpuff

(bottom) prior to behavioural training. **d**, auROC plots showing the firing rate changes in response to an expected (top) or unexpected (bottom) reward. **e**, auROC plots showing the firing rate changes in response to an expected (top) or unexpected (bottom) airpuff. **f**, auROC plots showing the firing rate changes in response to receiving an expected airpuff (left) or having an expected airpuff omitted (right). **g**, Histogram of difference in firing rate between airpuff omission and airpuff (filled bars,  $P < 0.05$ ,  $t$  test). Values are represented using auROC. **h**, auROC plots showing the firing rate changes in response to receiving an expected reward (left) or having an expected reward omitted (right). **i**, Histogram of difference in firing rate between reward omission and reward (filled bars,  $P < 0.05$ ,  $t$  test). Values are represented using auROC.

ACCELERATED ARTICLE PREVIEW



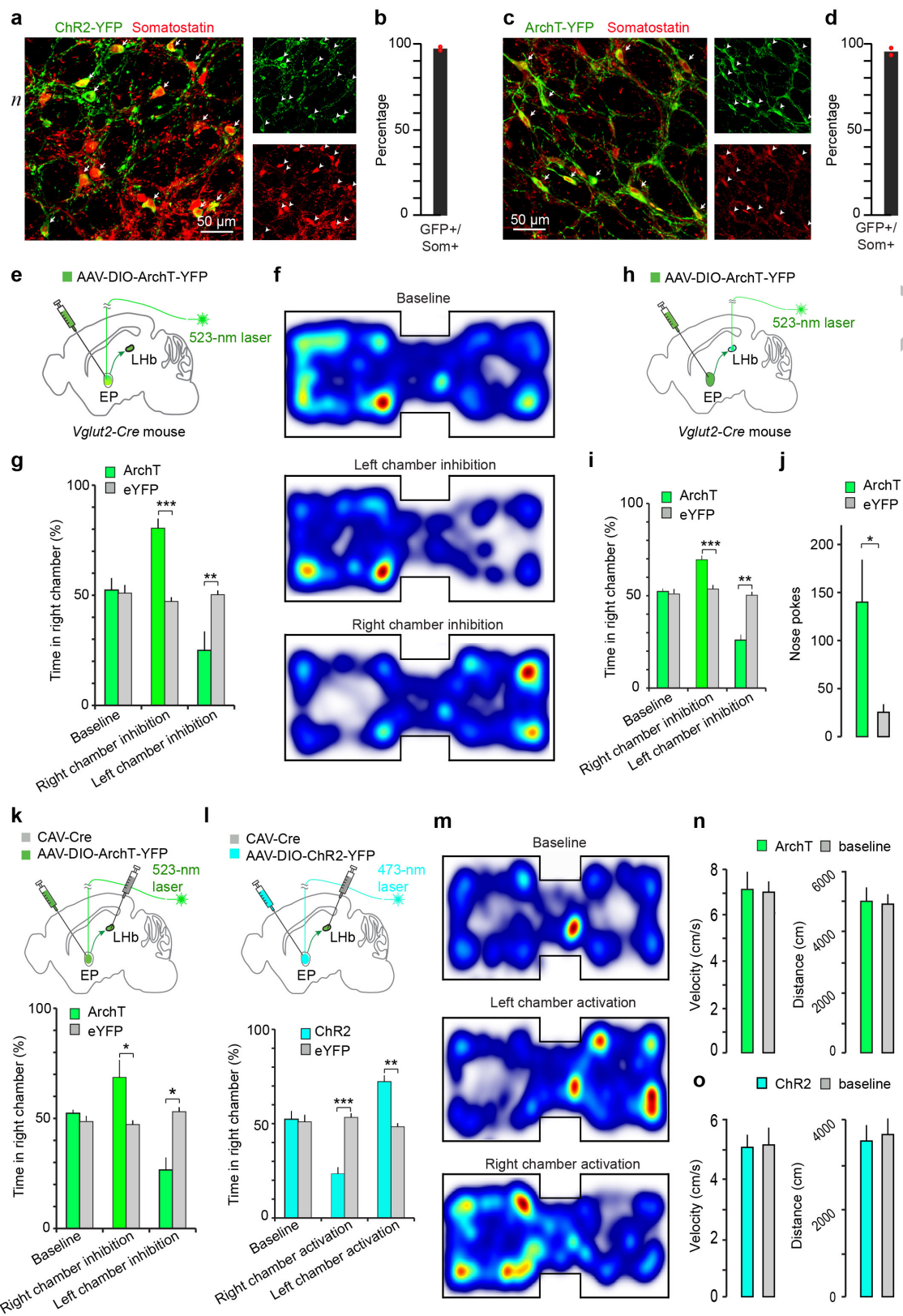


Extended Data Figure 4 | See next page for caption.

**Extended Data Figure 4 | Optic fiber implantation locations.**

**a**, A schematic of the experimental approach used for Arch-mediated inhibition of GPh neurons. **b**, A photomicrograph showing the location of optic fibre placement and ArchT-GFP<sup>+</sup> GPh neurons within the EP. **c**, Schematics showing the location of the optic fibre placements ( $n = 5$ ). **d**, A schematic of the experimental approach used for Arch-mediated inhibition of the GPh-LHb projection. **e**, A photomicrograph showing the location of optic fibre placement and ArchT-GFP<sup>+</sup> axon fibers within the LHb. **f**, Schematics showing the location of the optic fibre placements ( $n = 7$ ). **g**, A schematic of the experimental approach used for Arch-mediated inhibition of the GPh, which was targeted retrogradely by injection of the LHb with CAV2-Cre. **h**, A photomicrograph showing the location of the optic fibre placement and ArchT-GFP<sup>+</sup> neurons in the EP.

**i**, Schematics showing the location of the optic fibre placements ( $n = 5$ ). **j**, A schematic of the experimental approach used for ChR2-mediated excitation of the GPh, which was targeted retrogradely by injection of the LHb with CAV2-Cre. **k**, A photomicrograph showing the location of the optic fibre placement and ChR2-GFP<sup>+</sup> neurons in the EP. **l**, Schematics showing the location of the optic fibre placements ( $n = 5$ ). **m**, Schematic of the experimental approach used for ChR2-mediated activation of the GPh-LHb projection. **n**, A photomicrograph showing the optic fibre placement and ChR2-YFP<sup>+</sup> axon fibres in the LHb. **o**, Schematics showing the location of the optic fibre placements ( $n = 6$ ). Diagrams in **a**, **c**, **d**, **f**, **g**, **i**, **j**, **l**, **m** and **o** were modified from the Allen Mouse Brain Atlas, Allen Institute for Brain Science; available from <http://mouse.brain-map.org/>.

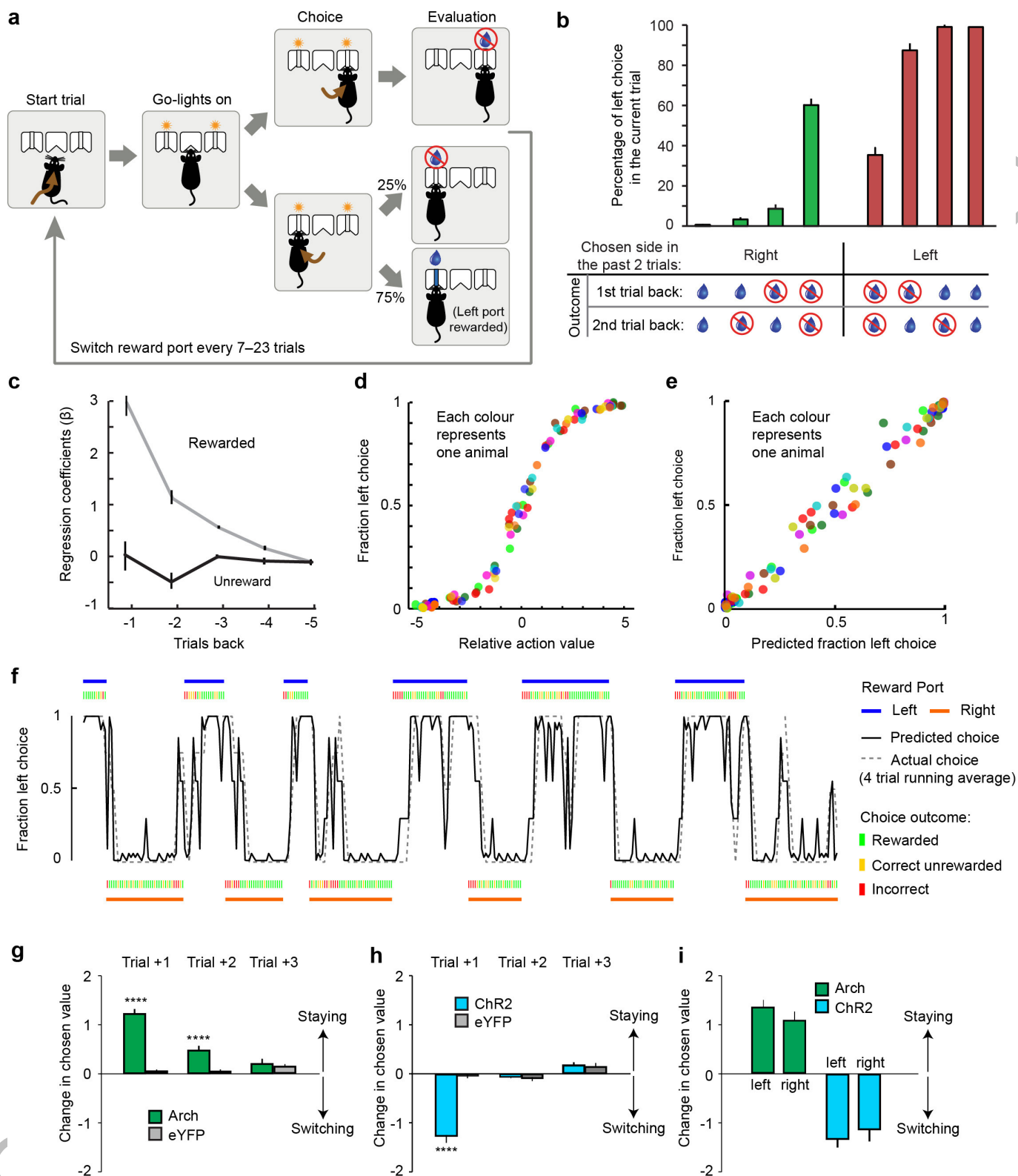


Extended Data Figure 5 | See next page for caption.



**Extended Data Figure 5 | Optogenetic inhibition of the GPh drives reward-related behaviours.** **a**, Confocal images from a *Som-cre;Ai14* mouse, showing the overlap in expression of ChR2-YFP and tdTomato (indicating Som+ neurons) in GPh neurons. **b**, Quantification of the percentage of ChR2-YFP+ neurons that expressed tdTomato ( $n = 2$ ). **c**, Confocal images from a *Som-cre;Ai14* mouse, showing the overlap in expression of ArchT-YFP and tdTomato in GPh neurons. **d**, Quantification of the percentage of ArchT-YFP+ neurons that expressed tdTomato ( $n = 2$ ). **e**, Schematic of the experimental approach used for ArchT-mediated inhibition of GPh neurons. **f**, Heatmaps for the activity of a representative mouse at baseline (top), or during optogenetic inhibition of the GPh in either the left (middle) or right (bottom) chamber. **g**, GPh<sup>Arch</sup> mice ( $n = 5$ ), but not GPh<sup>eYFP</sup> mice ( $n = 5$ ), showed a significant place preference for the chamber paired with laser stimulation in the GPh ( $F_{(5,29)} = 14.95$ ,  $P < 0.0001$ , \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , two-way ANOVA followed by Tukey's test). **h**, Schematic of the experimental approach used for ArchT-mediated inhibition of GPh axon terminals in the LHb. **i**, GPh<sup>ArchT</sup> mice ( $n = 7$ ), but not GPh<sup>eYFP</sup> mice ( $n = 5$ ), showed a significant place preference for the chamber paired with laser stimulation in the GPh ( $F_{(5,35)} = 52.22$ ,  $P < 0.0001$ , \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , two-way ANOVA followed by Tukey's test). **j**, GPh<sup>Arch</sup> mice ( $n = 5$ ) made significantly more nose pokes than GPh<sup>eYFP</sup> mice ( $n = 5$ ) to obtain laser

stimulation in the GPh ( $T_{(8)} = 2.61$ , \* $P < 0.05$ ,  $t$  test). **k**, Schematic of the retrograde labelling approach used to target the GPh for ArchT-mediated optical inhibition (top). GPh<sup>CAV-Cre/Arch</sup> mice ( $n = 5$ ), but not GPh<sup>eYFP</sup> mice ( $n = 5$ ), showed a significant place preference for the chamber paired with laser stimulation in the GPh (bottom) ( $F_{(5,29)} = 5.98$ ,  $P < 0.01$ , \* $P < 0.05$ , two-way ANOVA followed by Tukey's test). **l**, Schematic of the retrograde labelling approach used to target the GPh for ChR2-mediated optical excitation (top). GPh<sup>CAV-Cre/ChR2</sup> mice ( $n = 5$ ), but not GPh<sup>eYFP</sup> mice ( $n = 5$ ), showed a significant place aversion for the chamber paired with laser stimulation in the GPh (bottom) ( $F_{(5,29)} = 26.50$ ,  $P < 0.0001$ , \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , two-way ANOVA followed by Tukey's test). **m**, Heatmaps for the activity of a representative mouse at baseline (top), or during optogenetic excitation of the GPh in either the left (middle) or right (bottom) chamber. **n**, Mice did not move faster (left) or further (right) during the Arch stimulation sessions when compared to their baseline activity ( $T_{(32)} = 0.15$ ,  $P > 0.05$ ;  $T_{(32)} = 0.16$ ,  $P > 0.05$ ;  $t$  test,  $n = 17$ ). **o**, Mice did not move faster (left) or further (right) during the ChR2 stimulation sessions when compared to their baseline activity ( $T_{(8)} = 0.12$ ,  $P > 0.05$ ;  $T_{(8)} = 0.26$ ,  $P > 0.05$ ;  $t$  test,  $n = 5$ ). All data are presented as mean  $\pm$  s.e.m. Diagrams in **e**, **h**, **k**, and **i** were modified from the Allen Mouse Brain Atlas, Allen Institute for Brain Science; available from <http://mouse.brain-map.org/>.

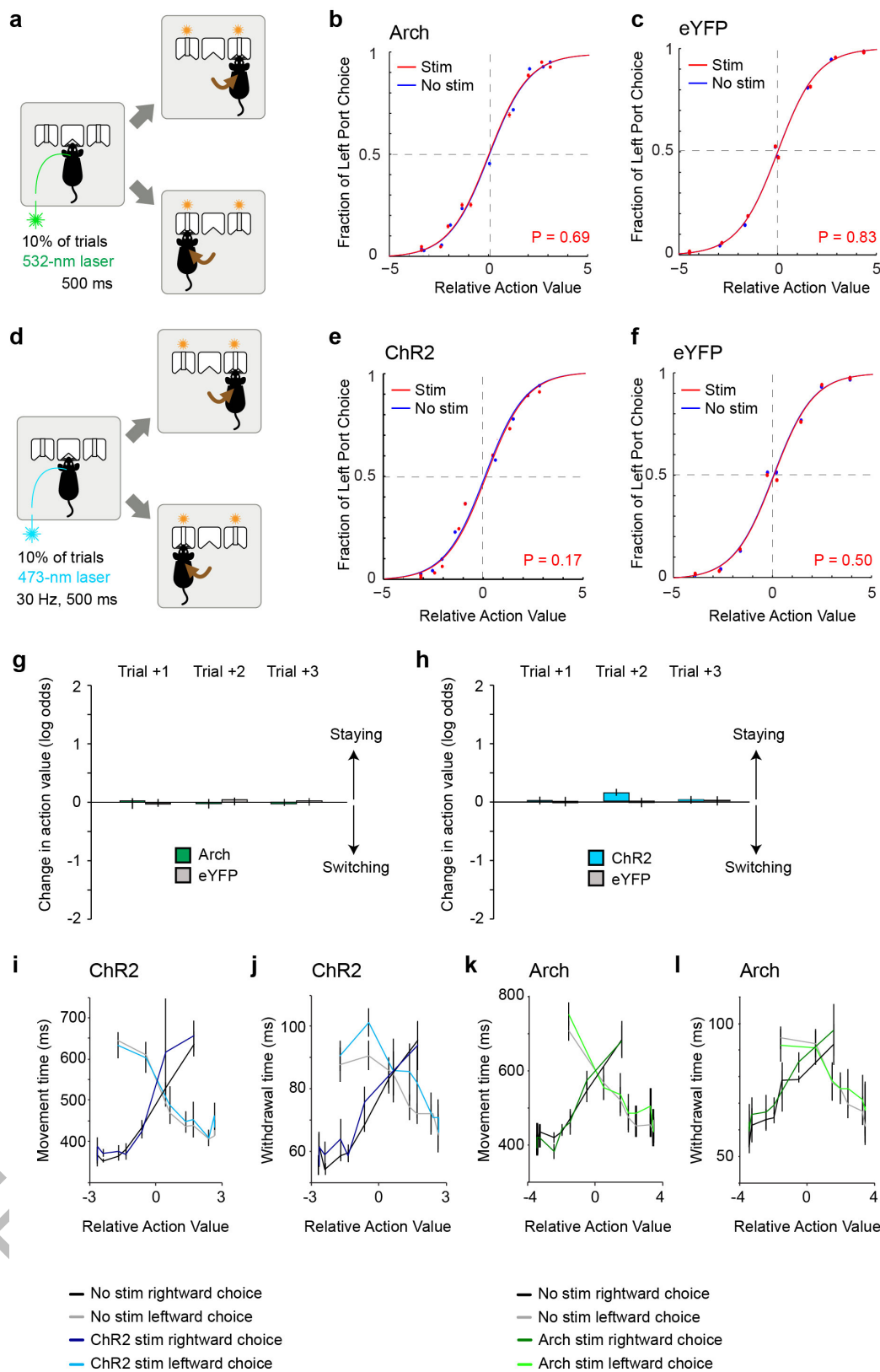


Extended Data Figure 6 | See next page for caption.

**Extended Data Figure 6 | A probabilistic switching task for studying action evaluation.** **a**, Schematic of the task. **b**, The probability of choosing the left port by one mouse for reward history in which consecutive choices to either the right or the left port were made during the previous two trials. **c**, The contribution of rewarded and unrewarded outcomes in the previous 5 trials – represented by regression coefficients  $\beta^{Reward}$  and  $\beta^{No\ Reward}$ , respectively – to choices in the current trial ( $n = 10$  mice,  $4685 \pm 786$  trials per mouse). **d**, The fraction of left port choice for 10 mice plotted against the relative action value (sum of the regression coefficients from the previous two trials). Data from each mouse was grouped into 10 bins and represented by a distinct colour. **e**, The actual probability of choosing the left port plotted against the probability of choosing the left port predicted by the logistic regression model. **f**, Example data from one session showing

12 trial blocks. Blue bars represent left reward blocks (top); orange bars indicate right reward blocks (bottom). Green, orange, and red ticks respectively represent whether a particular trial was a correct rewarded trial, a correct unrewarded trial, or an incorrect trial. The grey dashed line represents a four-trial running average of the mouse's probability of choosing the left port, and the black line indicates the probability of choosing the left port predicted by the logistic regression model. **g & h**, Change in chosen value one to three trials after optogenetic inhibition of the GPh (**g**), or activation of the GPh-LHb pathway (**h**). **i**, Changes in chosen value one trial after optogenetic activation or inhibition at the left or right reward port. In **g** and **h**, \*\*\*\* $P < 0.0001$ ,  $t$  test. In **b**, **c** and **g–i**, data are represented as mean  $\pm$  s.e.m.





Extended Data Figure 7 | See next page for caption.

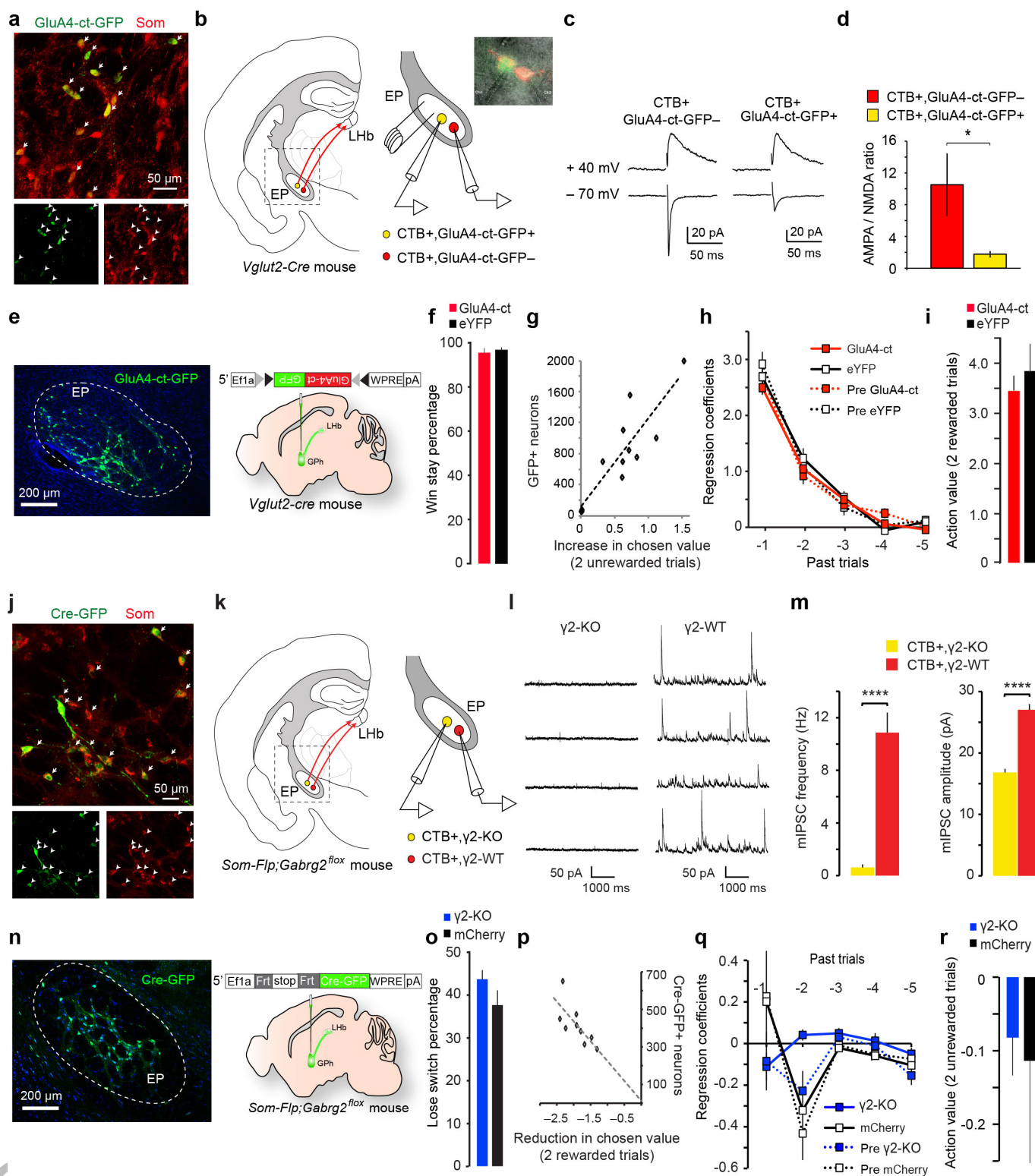
**Extended Data Figure 7 | Optogenetic inhibition or activation of the GPh-LHb pathway does not influence action selection.** **a**, A schematic of optogenetic inhibition of the GPh at the point of action selection.

**b**, Data points indicate the probability of left port choice as a function of action value for the trials in which the photo-stimulation was delivered at the center port ("stim") or was not delivered ("no stim"). Lines indicate the fit by the logistic regression model on the pooled data for each of the two conditions ( $n = 5$  mice, 15,411 trials,  $3082 \pm 1063$  trials per mouse).

**c**, Similar to **b**, except that control mice with eYFP-expressing GPh neurons were used ( $n = 6$  mice, 56,241 trials,  $9373 \pm 596$  trials per mouse).

**d & e**, Similar to **a** and **b**, except that optogenetic activation of the GPh-LHb projection was applied at the point of action selection ( $n = 6$  mice, 41,557 trials,  $8311 \pm 2565$  trials per mouse). **f**, Similar to **e**, except that control mice with eYFP-expressing GPh neurons were used ( $n = 6$  mice, 72,423 trials,  $12070 \pm 1673$  trials per mouse). **g & h**, The changes in action value in response to optogenetic stimulation of the GPh-LHb pathway one to three trials after the photo-stimulation, for mice in which GPh neurons expressed Arch ( $n = 5$ ) or eYFP ( $n = 6$ ) (**g**), or Chr2 ( $n = 6$ ) or eYFP ( $n = 6$ ) (**h**). In **b**, **c**, **e**, and **f**,  $P$  values reported for  $t$  tests:  $H_0$ :

$\beta_{\text{stim}} = 0$ . **i-l**, Graphs showing the average withdrawal (calculated as the time from center port entry to exit) and movement (calculated as the time from center port exit to the poke at the chosen port) time for trials with or without light stimulation. Both withdrawal time and movement time were shorter when the action value associated with the chosen action was higher. Neither activation of GPh neurons with Chr2 ( $n = 6$  mice) (**i & j**) (movement time for leftward choices, Chr2 stimulated trials ("Chr2") vs. unstimulated trials ("no stim"),  $F_{(1,8)} = 0.174$ ,  $P > 0.05$ ; rightward choices, Chr2 vs. no stim,  $F_{(1,8)} = 1.352$ ,  $P > 0.05$ ; withdrawal time preceding leftward choices, Chr2 vs. no stim,  $F_{(1,8)} = 0.667$ ,  $P > 0.05$ ; preceding rightward choices, Chr2 vs. no stim,  $F_{(1,8)} = 0.599$ ,  $P > 0.05$ ; two way ANOVA), nor inhibition of these neurons with Arch ( $n = 5$  mice) (**k & l**) (movement time for leftward choices, Arch stimulated trials ("Arch") vs. unstimulated trials ("no stim"),  $F_{(1,8)} = 0.105$ ,  $P > 0.05$ ; rightward choices, Arch vs. no stim,  $F_{(1,8)} = 0.023$ ,  $P > 0.05$ ; withdrawal time preceding leftward choices, Arch vs. no stim,  $F_{(1,8)} = 0.821$ ,  $P > 0.05$ ; preceding rightward choices, Arch vs. no stim,  $F_{(1,8)} = 0.459$ ,  $P > 0.05$ ; two way ANOVA) had any significant effect on the ongoing behaviour. Data in **g-l** are presented as mean  $\pm$  s.e.m.

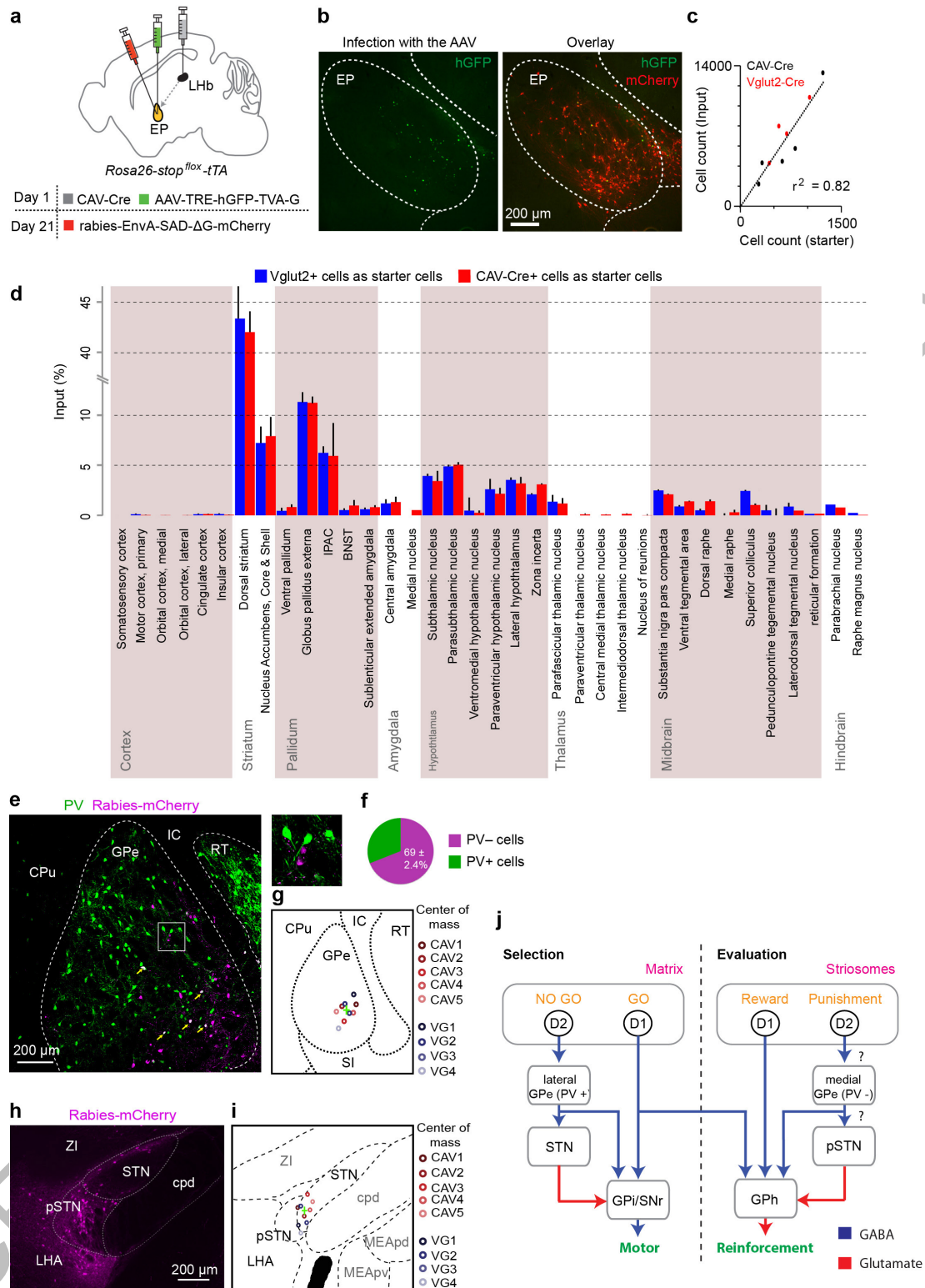


Extended Data Figure 8 | See next page for caption.



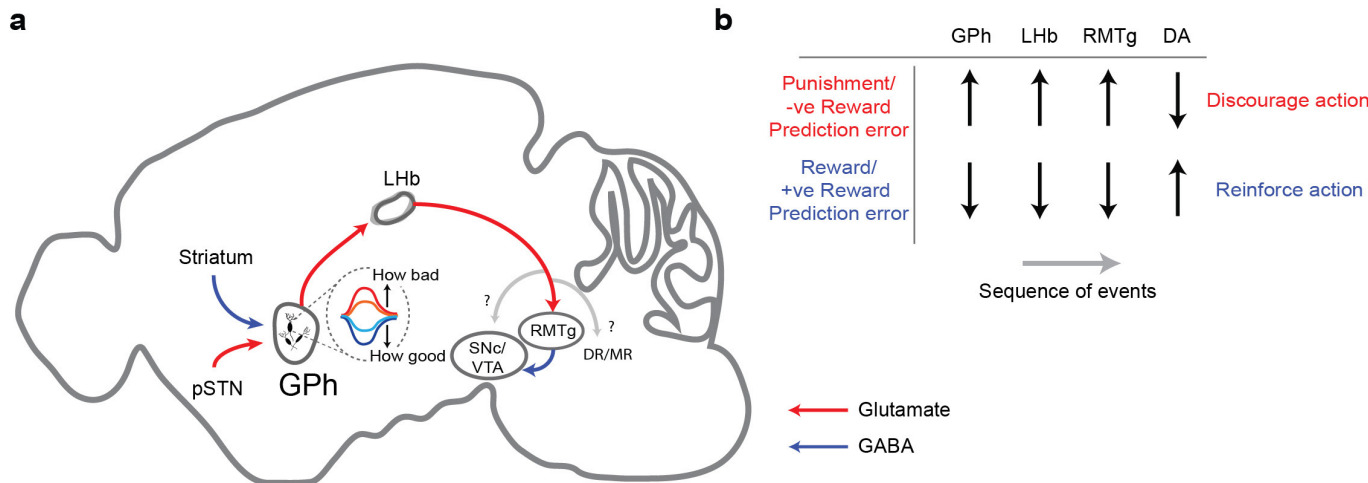
**Extended Data Figure 8 | Weakening of excitatory or inhibitory synapses onto GPh neurons and its effects on the sensitivity to negative or positive feedback.** **a**, Confocal images from a *Som-cre;Ai14* mouse, showing the overlap in expression of GluA4-ct-GFP (delivered by injecting the EP with the AAV-DIO-GluA4-ct-GFP) and tdTomato (indicating the expression of Som) in GPh neurons.  $97.86 \pm 2.9\%$  of GluA4-ct-GFP+ neurons expressed tdTomato ( $n = 2$  mice). **b**, Schematics of the experimental approach. CTB-594 was injected into the LHB to label GPh neurons in the EP. On the right is an enlarged graph of the boxed area in the cartoon on the left. Inset is a photomicrograph showing simultaneous recording of a CTB<sup>+</sup>/GluA4-ct<sup>+</sup> GPh neuron and a nearby CTB<sup>+</sup>/GluA4-ct<sup>-</sup> GPh neuron. **c**, EPSC traces recorded from the two neurons shown in **b**. **d**, Quantification of the ratio between AMPA receptor-mediated EPSC amplitude and NMDA receptor-mediated EPSC amplitude (AMPA/NMDA ratio) for the two populations of GPh neurons (CTB<sup>+</sup>/GluA4-ct<sup>+</sup>,  $n = 6$  cells; CTB<sup>+</sup>/GluA4-ct<sup>-</sup>,  $n = 8$  cells;  $n = 3$  mice;  $T_{(12)} = -1.89$ ,  $*P < 0.05$ ,  $t$  test). **e**, A representative image showing the expression of GluA4-ct-GFP (delivered by injecting the EP of a *Vglut2-Cre* mouse with the AAV-DIO-GluA4-ct-GFP) in GPh neurons (left) and a schematic of the approach (right). **f**, The win-stay percentage in these mice (GPh<sup>GluA4-ct</sup>,  $94.17 \pm 1.02\%$ ; GPh<sup>eYFP</sup>,  $95.82 \pm 0.51\%$ ;  $P > 0.05$ ,  $t$  test). **g**, For animals ( $n = 10$  mice) used in Fig. 4a-e, the number of GPh neurons that were infected with the GluA4-ct-GFP virus correlated with the change in animal behaviour in the switching task, measured as an increase in action value following two consecutive unrewarded trials ( $R^2 = 0.72$ ,  $P < 0.05$  by a linear regression). **h**, Contributions of rewarded outcomes over the past five trials, as reflected by their regression coefficients, to the current choice. GPh<sup>GluA4-ct</sup> mice were not significantly different from control mice or their pre-surgery condition (first two trials back x groups,  $F_{(3,33)} = 0.5412$ ,  $P > 0.05$ ; two-way ANOVA,  $n = 10$  GPh<sup>GluA4-ct</sup> mice and  $n = 7$  control mice). **i**, The action value following two sequentially rewarded trials was not significantly different between

GPh<sup>GluA4-ct</sup> mice and GPh<sup>eYFP</sup> mice ( $P > 0.05$ ,  $t$  test). **j**, Confocal images from a *Som-flp* mouse, showing the overlap in expression of Cre-GFP (delivered by injecting the EP with the AAV-FSF-GFP-Cre) and somatostatin, recognized through antibody labelling.  $96.25 \pm 2.3\%$  of Cre-GFP+ neurons expressed somatostatin ( $n = 2$  mice). **k**, Schematics of the experimental approach. CTB-594 was injected into the LHB to label GPh neurons in the EP. On the right is an enlarged graph of the boxed area in the cartoon on the left. **l**, Sample miniature IPSC (mIPSC) traces recorded from a GPh neuron that expressed Cre-GFP – and thus had  $\gamma 2$  ablated ( $\gamma 2$ -KO) – and a control GPh neuron that did not express the Cre-GFP ( $\gamma 2$ -WT). **m**, Quantification of the frequency (left) and amplitude (right) of mIPSCs recorded from the two groups of GPh neurons ( $\gamma 2$ -KO,  $n = 7$  cells;  $\gamma 2$ -WT,  $n = 10$  cells;  $n = 3$  mice; frequency,  $T_{(15)} = 5.51$ ,  $****P < 0.0001$ ; amplitude,  $T_{(15)} = 8.19$ ,  $****P < 0.0001$ ;  $t$  test). **n**, A representative image showing the expression of Cre-GFP (delivered by injecting the EP of a *Som-Flp;Gabra2<sup>fllox</sup>* mouse with the AAV-FSF-GFP-Cre) in GPh neurons (left) and a schematic of the approach (right). **o**, The lose-switch percentage in these mice ( $P > 0.05$ ,  $t$  test). **p**, For animals ( $n = 9$ ) used in Fig. 4f-j, the number of GPh neurons that were infected with the Cre-GFP virus correlated with the change in animal behaviour in the switching task, measured as a reduction in action value following two consecutive rewarded trials ( $R^2 = 0.53$ ,  $P < 0.05$  by a linear regression). **q**, The negative regression coefficients associated with the past five trials were not significantly different between GPh <sup>$\gamma 2$ -KO</sup> mice and control mice either before or after surgery (first two trials back x groups,  $F_{(3,35)} = 0.9072$ ,  $P > 0.05$ ,  $n = 9$  GPh <sup>$\gamma 2$ -KO</sup> mice and  $n = 9$  control mice). **r**, The action value following two sequentially unrewarded trials was not significantly different between GPh <sup>$\gamma 2$ -KO</sup> mice and GPh<sup>mCherry</sup> mice ( $P > 0.05$ ,  $t$  test). All data are represented as mean  $\pm$  s.e.m. Diagrams in **b**, **e**, **k** and **n** were modified from the Allen Mouse Brain Atlas, Allen Institute for Brain Science; available from <http://mouse.brain-map.org/>.



**Extended Data Figure 9 | Monosynaptic inputs onto the GPe and a schematic of the circuitry for reinforcement learning.** **a**, Schematics of experimental design. The GPe neurons in the EP were targeted using either *Vglut2-Cre;Rosa26-stop<sup>fllox</sup>-tTA* mice or by injecting the LHb of *Rosa26-stop<sup>fllox</sup>-tTA* mice with the retrograde CAV2-Cre. **b**, Images showing the starter cell location in the EP. **c**, Relationship between the number of starter and input neurons. **d**, Graph showing the fraction of monosynaptically labelled neurons in each brain region that projects to the GPe ( $n = 9$  mice). **e**, Confocal images of the rabies virus and parvalbumin (PV) labelled neurons in the GPe. Only a small fraction of the virally labelled GPe cells expressed PV (arrows). On the right is a high

magnification image of the boxed area in the GPe. **f**, Quantification of the fraction of rabies virus labelled GPe neurons that expressed PV ( $n = 3$  mice). **g**, Center of mass analysis for all GPe labelled neurons ( $n = 9$  mice). **h**, A confocal image of the parasubthalamic nucleus (pSTN) showing monosynaptically labelled neurons. **i**, Center of mass analysis for all pSTN labelled neurons ( $n = 9$  mice). **j**, A schematic showing the proposed selection and evaluation circuits within the basal ganglia. Question marks indicate elements of the proposed circuit that remain to be tested experimentally. Diagrams in **a**, **g** and **i** were modified from the Allen Mouse Brain Atlas, Allen Institute for Brain Science; available from <http://mouse.brain-map.org/>.



**Extended Data Figure 10 | The proposed function of the basal ganglia and midbrain evaluation circuits.** **a**, schematic showing the activity of GPh neurons and the downstream circuitry controlling the midbrain dopaminergic system. **b**, Proposed sequence of events by which GPh activity may influence the firing rate in downstream structures. Upward arrows indicate an increase in firing; downward arrows indicate a decrease in firing. RMTg, Rostromedial tegmental nucleus; SNc, Substantia nigra

pars compacta; VTA, ventral tegmental area; DA, dopamine. DR, dorsal raphe; MR, median raphe. ? indicates that alternative circuits downstream of the LHb, including the serotonergic raphe nuclei, may constitute other key pathways that also process the GPh-LHb prediction error signals that we demonstrate in this study. Diagram in **a** was modified from the Allen Mouse Brain Atlas, Allen Institute for Brain Science; available from <http://mouse.brain-map.org/>.



# Fatty acid synthesis configures the plasma membrane for inflammation in diabetes

Xiaochao Wei<sup>1</sup>, Haowei Song<sup>1</sup>, Li Yin<sup>1</sup>, Michael G. Rizzo Jr<sup>1</sup>, Rohini Sidhu<sup>2</sup>, Douglas F. Covey<sup>3</sup>, Daniel S. Ory<sup>2</sup> & Clay F. Semenkovich<sup>1,4</sup>

**Dietary fat promotes pathological insulin resistance through chronic inflammation<sup>1–3</sup>. The inactivation of inflammatory proteins produced by macrophages improves diet-induced diabetes<sup>4</sup>, but how nutrient-dense diets induce diabetes is unknown<sup>5</sup>. Membrane lipids affect the innate immune response<sup>6</sup>, which requires domains<sup>7</sup> that influence high-fat-diet-induced chronic inflammation<sup>8,9</sup> and alter cell function based on phospholipid composition<sup>10</sup>. Endogenous fatty acid synthesis, mediated by fatty acid synthase (FAS)<sup>11</sup>, affects membrane composition. Here we show that macrophage FAS is indispensable for diet-induced inflammation. Deleting *Fasn* in macrophages prevents diet-induced insulin resistance, recruitment of macrophages to adipose tissue and chronic inflammation in mice. We found that FAS deficiency alters membrane order and composition, impairing the retention of plasma membrane cholesterol and disrupting Rho GTPase trafficking—a process required for cell adhesion, migration and activation. Expression of a constitutively active Rho GTPase, however, restored inflammatory signalling. Exogenous palmitate was partitioned to different pools from endogenous lipids and did not rescue inflammatory signalling. However, exogenous cholesterol, as well as other planar sterols, did rescue signalling, with cholesterol restoring FAS-induced perturbations in membrane order. Our results show that the production of endogenous fat in macrophages is necessary for the development of exogenous-fat-induced insulin resistance through the creation of a receptive environment at the plasma membrane for the assembly of cholesterol-dependent signalling networks.**

LysM–FAS mice (mice with a LysM–Cre-induced myeloid cell FAS deficiency) have normal glucose tolerance when fed chow, but improved glucose tolerance on a high-fat diet (HFD), lower blood glucose in response to insulin, and lower insulin levels when compared to controls, despite exhibiting no difference in body composition or weight (Fig. 1a–d). Insulin-stimulated phosphorylation of Akt was increased in the adipose tissue and livers of LysM–FAS mice (Fig. 1e, f), indicating increased insulin sensitivity. There were fewer crown-like structures (Fig. 1g, h) and total (Fig. 1i), as well as pro-inflammatory (Extended Data Fig. 1a), macrophages in the visceral fat of LysM–FAS mice than in controls. Levels of inflammatory gene expression (Fig. 1j) and phosphorylated c-Jun N-terminal Kinase (pJNK; Fig. 1k), which promotes obesity-associated insulin resistance, were decreased in the adipose tissue of HFD-fed LysM–FAS mice when compared to control animals. Steatosis (Fig. 1l–n) and inflammatory gene expression (Fig. 1o) were also decreased in the livers of HFD-fed LysM–FAS mice. These results suggest that macrophage FAS promotes diet-induced insulin resistance.

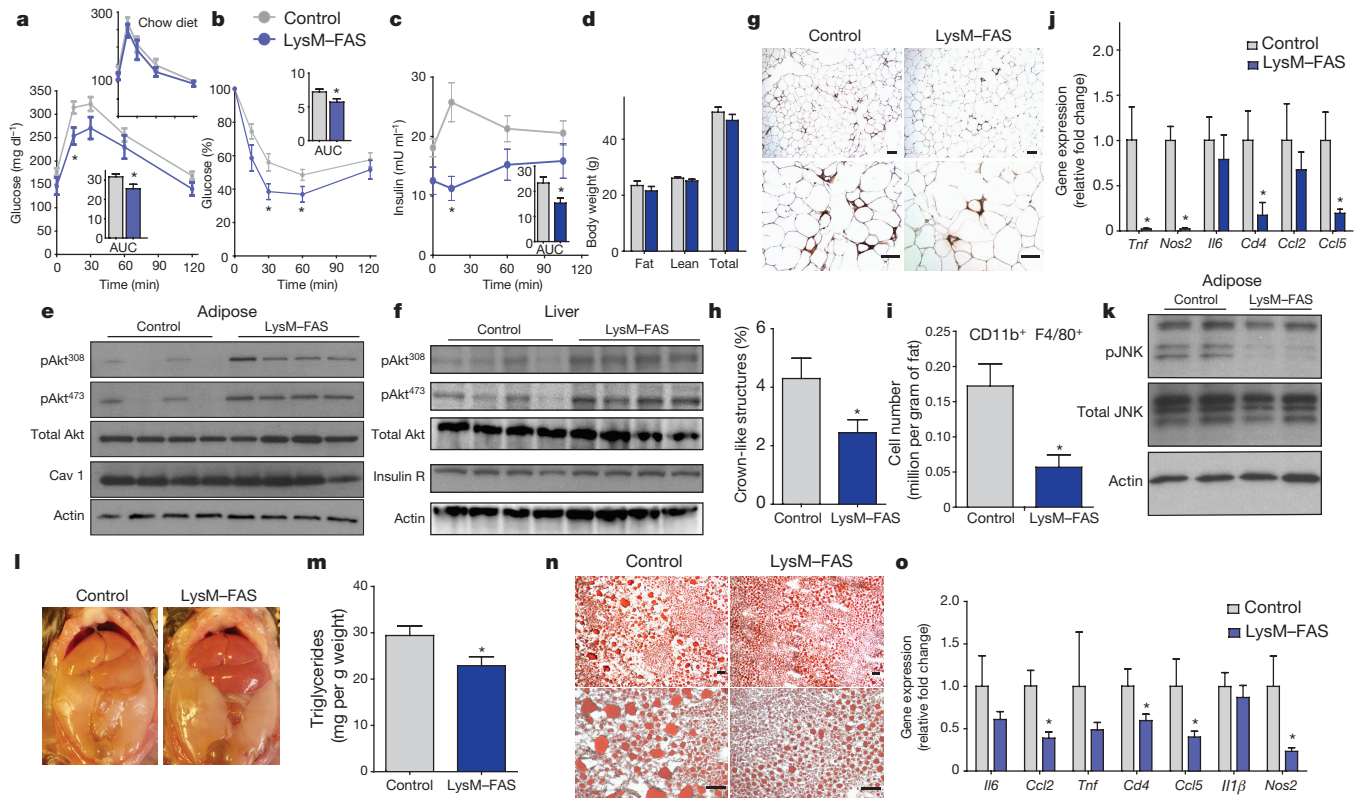
Levels of FAS increased when bone-marrow-derived macrophages from control mice or cells of the mouse macrophage-like cell line RAW 264.7 were exposed to high-dose palmitate or lipopolysaccharide

(LPS; Extended Data Fig. 1b–e), indicating that endogenous fatty acid synthesis is associated with macrophage activation. Peritoneal macrophages from LysM–FAS mice exhibited decreased levels of pJNK and inflammatory cytokine generation in response to LPS (Fig. 2a, b) or palmitate (Fig. 2c, d) when compared to controls. Pharmacological inhibition of FAS enzyme activity decreased LPS-induced JNK phosphorylation (Extended Data Fig. 1f), with knockdown of FAS in RAW 264.7 cells decreasing JNK phosphorylation and inflammatory cytokine generation (Extended Data Fig. 1g–k).

Tie2–FAS mice (mice with a Tie2–Cre-induced endothelial and haematopoietic cell FAS deficiency) display defective angiogenesis but normal glucose levels on a chow diet<sup>12</sup>. Tie2–FAS mice and wild-type mice infused with bone marrow from Tie2–FAS mice were, when compared to respective controls, protected from diet-induced insulin resistance and inflammation (Extended Data Figs 2–4). This indicates that FAS deficiency, as seen in different Cre mice and with varied genetic and chemical approaches in cultured cells, decreases macrophage activation.

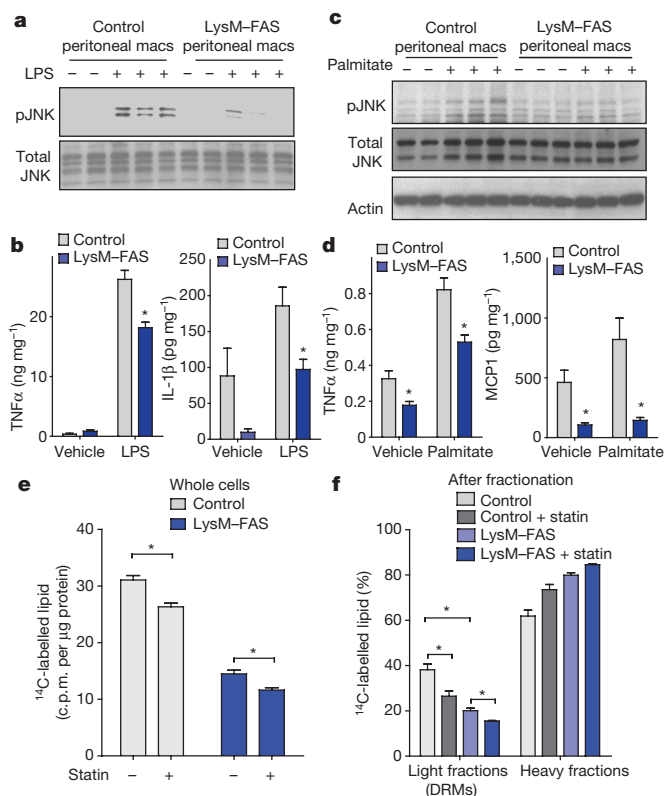
The incubation of macrophages with radiolabelled <sup>14</sup>C-acetate helped to demonstrate the distinct effects of inhibiting fatty acid and cholesterol synthesis on whole-cell accumulation of labelled lipids (Fig. 2e), with effects mostly reflected in labile detergent-resistant microdomains (DRMs) (Fig. 2f). This suggests that FAS-dependent lipids and newly synthesized sterols are channelled to DRMs. Levels of DRM-associated glycerophospholipids were decreased in FAS-deficient macrophages but there was a minimal effect on whole-cell membranes (Extended Data Fig. 5), suggesting that FAS deficiency alters microdomain phospholipids while preserving whole-membrane lipid composition. Proteomic analysis<sup>13</sup> of DRMs from FAS-replete macrophages from control mice and from FAS-deficient macrophages from LysM–Cre and Tie2–Cre mice showed that 534 of 794 proteins were reduced by more than 40% in FAS-deficient DRMs (Extended Data Fig. 6a, b and Supplementary Table 1). In whole membranes, however, only 17 of 681 proteins were reduced by more than 40% in FAS-deficient macrophages (Extended Data Fig. 6c, d and Supplementary Table 2). LysM–FAS and Tie2–FAS models showed coordinated suppression of the same proteins in DRMs but little effect on whole-membrane protein content (Extended Data Fig. 6e). FAS affected proteins involved both in phagocytosis (Extended Data Fig. 6f, top) and in responses to pathogens (Extended Data Fig. 6f, middle), functions known to be DRM-dependent<sup>14</sup>. Proteins that mediate macrophage inflammation also appeared to require FAS for DRM localization (Extended Data Fig. 6f, g). In both FAS-deficient models (Extended Data Fig. 6h), nearly all of these proteins showed decreased abundance in DRMs but not in whole membranes, results that were confirmed by western blots (Extended Data Fig. 6i) including control proteins that were unaffected by the lack of FAS such as the archetypal non-DRM protein transferrin receptor (TfR) and the DRM proteins flotillin-1 and Lyn.

<sup>1</sup>Division of Endocrinology, Metabolism and Lipid Research, Washington University School of Medicine, St. Louis, Missouri 63110, USA. <sup>2</sup>Diabetic Cardiovascular Disease Center, Washington University School of Medicine, St. Louis, Missouri 63110, USA. <sup>3</sup>Department of Developmental Biology, Washington University School of Medicine, St. Louis, Missouri 63110, USA. <sup>4</sup>Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, Missouri 63110, USA.



**Figure 1 | Macrophage FAS ablation ameliorates diet-induced insulin resistance and inflammation in mice.** **a–c**, Glucose tolerance test (**a**), insulin tolerance test (**b**), and glucose-stimulated insulin secretion (**c**) in LysM-FAS mice fed a HFD for 3 months (inset in **a** shows results for chow diet;  $n = 10$  control and 7 LysM-FAS). AUC, area under the curve. **d**, Body composition for the mice in **a–c**. **e**, **f**, Insulin-stimulated Akt phosphorylation in adipose tissue (**e**) and liver (**f**) of control and LysM-FAS mice. **g–i**, Photomicrographs (**g**), crown-like structures (**h**), and

macrophages (**i**) in visceral adipose tissue of HFD-fed mice ( $n = 4$  per group). **j**, **k**, Inflammatory gene expression (**j**) and JNK phosphorylation (**k**) in adipose tissue of control ( $n = 7$ ) and LysM-FAS ( $n = 6$ ) mice. **l**, **m**, Gross appearance (**l**) and triglyceride content (**m**) of liver from control ( $n = 6$ ) and LysM-FAS mice ( $n = 8$ ). **n**, **o**, Oil red O staining (**n**) and inflammatory gene expression (**o**) in livers from control ( $n = 6$ ) and LysM-FAS mice ( $n = 8$ ). \* $P < 0.05$ ; data are mean  $\pm$  s.e.m. (see Methods for details of the statistical test used). Scales bars, 50  $\mu$ m (**g**), 100  $\mu$ m (**n**).



**Figure 2 | Macrophage FAS deficiency attenuates cell-autonomous inflammation and alters detergent-resistant microdomains.** **a**, **b**, JNK phosphorylation blot (**a**) and quantification of cytokine stimulation (**b**) by LPS in peritoneal macrophages (macs). **c**, **d**, JNK phosphorylation blot (**c**) and quantification of cytokine stimulation (**d**) by palmitate in peritoneal macrophages. **e**, **f**, <sup>14</sup>C radioactivity in whole-cell lipid extracts from control and FAS-knockout macrophages in the absence or presence of simvastatin (**e**; 10  $\mu$ M, 24 h), and per cent radioactivity in lipid extracts from light or heavy fractions (**f**;  $n = 4$  per condition). Enzyme-linked immunosorbent assays (ELISAs) (**b** and **d**) were performed in triplicate for each isolate and repeated three times using different mice. \* $P < 0.05$ ; data are mean  $\pm$  s.e.m. (see Methods for details of the statistical test used).

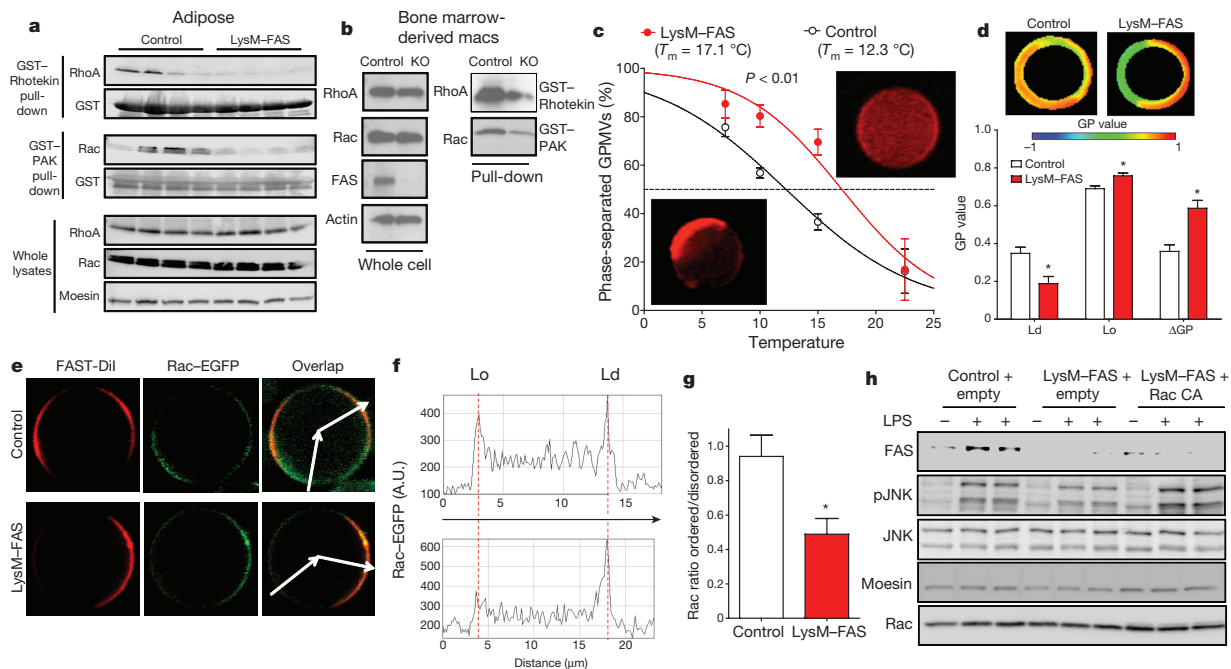
Cdc42, Rac, moesin, RhoA, myosin and actin, proteins that require dynamic trafficking through DRMs<sup>14–20</sup>, were decreased in FAS-deficient DRMs (Extended Data Fig. 6i). Activities of Rho family GTPases were reduced in the adipose tissue of LysM-FAS (Fig. 3a) and Tie2-FAS bone-marrow-transplanted (Extended Data Fig. 7a) mice fed on a HFD, as well as in FAS-deficient macrophages (Fig. 3b) and RAW 264.7 cells (Extended Data Fig. 7b). Rho GTPases affect cell motility: cell adhesion and migration were decreased in LysM-FAS macrophages (Extended Data Fig. 7c–e) and FAS-knockdown RAW 264.7 cells (Extended Data Fig. 7f–h) when compared to controls. These findings suggest that FAS is required for macrophage inflammatory signalling as it maintains a lipid environment that is conducive to DRM protein trafficking.

Since DRM isolation can induce artefacts, we studied giant plasma membrane vesicles (GPMVs)<sup>21</sup>, structures that exist at thermodynamic equilibrium. Miscibility transition temperature (the temperature at which 50% of GPMVs are phase-separated, as detected using the liquid-disordered domain dye DiI) was increased in GPMVs from LysM-FAS mice when compared to control macrophages (Fig. 3c), indicating a FAS-dependent difference in the separation of liquid-disordered and liquid-ordered domains. An increase in the difference in membrane order for distinct phases (as determined using dyes with packing-dependent emission) correlates with increased miscibility transition temperatures<sup>22</sup>. Using the phase-sensitive dye Di-4-ANEPPDHQ, we found that the difference in general polarization was greater in LysM-FAS mouse GPMVs than in controls (Fig. 3d), confirming that membrane order was altered in FAS-deficient cells. When LysM-FAS and control cells were transfected with EGFP-tagged Rac followed by GPMV preparation,

the Rac-GFP signal was present in the liquid-disordered domain but decreased in the liquid-ordered domain of FAS-deficient cells (Fig. 3e–g). Rac-GFP partitioning to the liquid-ordered domain was also diminished when GPMVs were prepared from FAS-deficient RAW 264.7 cells (Extended Data Fig. 8a, b). These results suggest that FAS deficiency alters membrane composition and thereby interferes with inflammatory signalling by limiting the assembly of signalling domains that contain Rho GTPases. Transfection of FAS-deficient cells with constitutively active Rac under conditions that did not increase total Rac content restored the activation of JNK by LPS in intact LysM-FAS cells (Fig. 3h). These results confirm that FAS-deficient cells retain the capacity to transmit inflammatory signals and suggest that FAS is important for configuring membrane composition to facilitate the recruitment of signalling molecules.

In pulse-labelling experiments, endogenous lipids (<sup>14</sup>C-acetate-labelled fatty acids and cholesterol) were rapidly chased from DRMs (Fig. 4a, left) but exogenous labelled palmitate accumulated in DRMs (Fig. 4a, right). These results confirm that exogenous palmitate may accumulate in DRMs<sup>9</sup>, but also indicate that endogenous and exogenous lipids access different pools. Since DRMs require endogenous lipids, exogenous palmitate should not rescue the FAS-deficient phenotype. Consistent with this, pretreatment with 50  $\mu$ M palmitate did not restore the JNK response to LPS (Fig. 4b) or high-dose (500  $\mu$ M) palmitate (Extended Data Fig. 9a). Exposure to chronic exogenous 50  $\mu$ M palmitate did not normalize the DRM Rho GTPase content of LysM-FAS cells and did not correct migration or adhesion defects (Extended Data Fig. 9b–d).

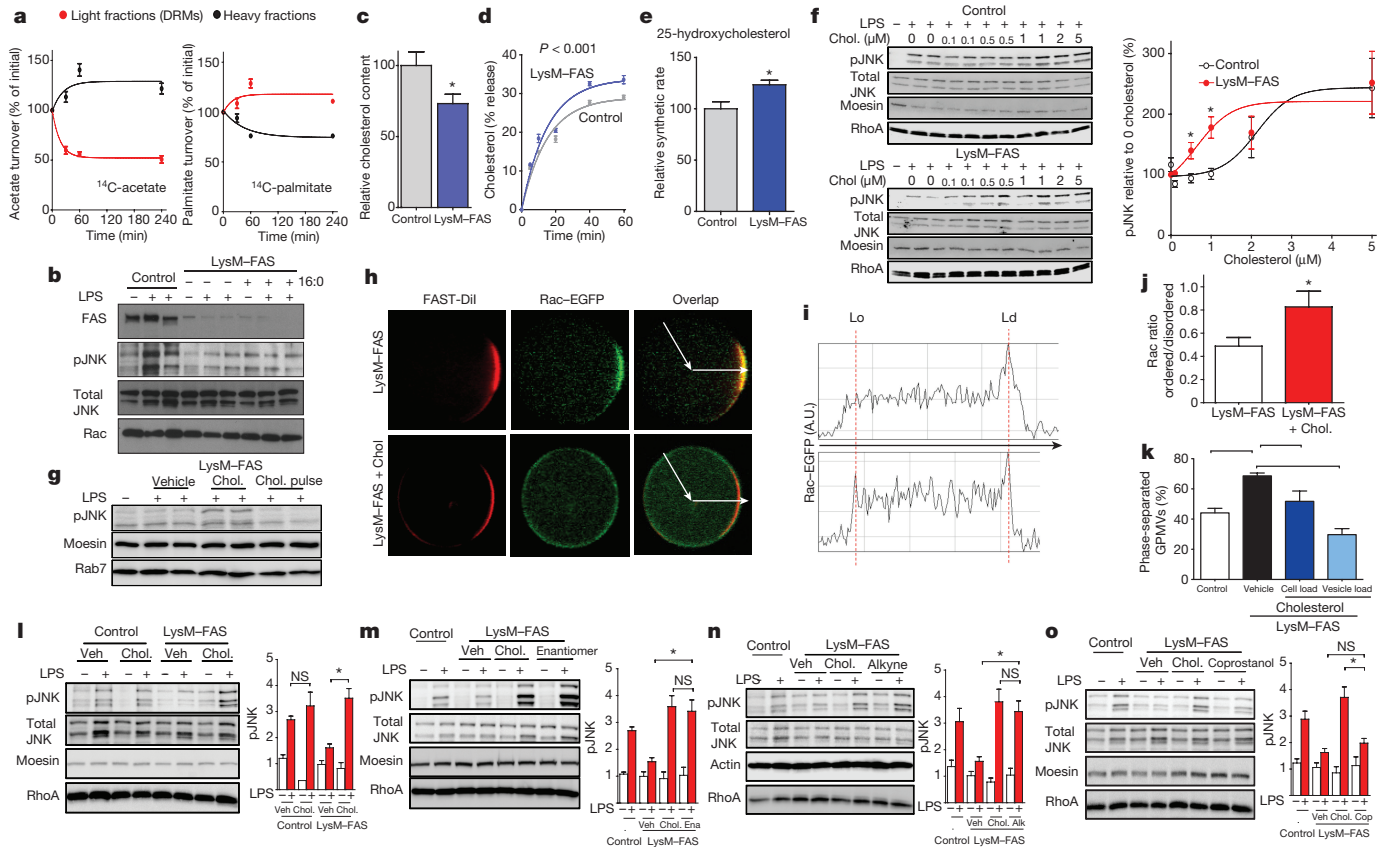
Cholesterol levels were decreased with FAS deficiency (Fig. 4c and Extended Data Fig. 9e, f). Compared to controls, cholesterol was more



**Figure 3 | FAS deficiency disrupts Rho GTPases and alters membrane order.** **a**, Western blots of activated RhoA (top), activated Rac (middle), and total RhoA and Rac (bottom) in adipose tissue of control and LysM-FAS mice with the indicated controls. GST, glutathione S-transferase. **b**, Total (left) or activated (right) Rho GTPases in control or FAS-knockout (KO) bone-marrow-derived macrophages. **c**, GPMVs prepared from bone-marrow-derived macrophages were labelled with the liquid-disordered fluorescent dye FAST-Dil and cooled in a step-wise manner followed by scoring vesicles for phase separation by confocal microscopy. Insets show phase-separated (lower left) and non-separated (upper right) vesicles. **d**, General polarization values were determined using the dye Di-4-ANEPPDHQ. Ld, liquid-disordered domain; Lo, liquid-ordered domain. **e**, Cells were transfected with a Rac-EGFP construct followed by

preparation of GPMVs and staining with the liquid-disordered domain dye FAST-Dil. Liquid-ordered and liquid-disordered domains for quantification are indicated by arrows. **f**, Rac-EGFP levels in liquid-ordered and liquid-disordered domains. **g**, Ratio of Rac partitioning in ordered to disordered domains in control and LysM-FAS cells. **h**, Western blots for FAS, pJNK and Rac in control cells and LysM-FAS cells transfected with an empty vector or a gene encoding constitutively active Rac (Rac CA), followed by determination of JNK phosphorylation with LPS stimulation. Data in **c** are pooled ( $n = 492$  total vesicles analysed for control and  $n = 453$  total vesicles analysed for LysM-FAS) from 3 different mouse pairs. Data in **d**, ( $n = 6$  per group) and **g**, ( $n = 5$  per group) are representative experiments. \* $P < 0.05$ , data are mean  $\pm$  s.e.m. (see Methods for details of the statistical test used).





**Figure 4 | Exogenous palmitate does not rescue inflammation, but exogenous cholesterol restores JNK signalling, membrane order, and Rho GTPase membrane distribution. a**, Pulse-chase turnover of  $^{14}\text{C}$ -acetate (left) or  $^{14}\text{C}$ -palmitate-labelled lipids in DRMs versus heavy fractions ( $n = 4$ ). **b**, Chronic incubation with  $50 \mu\text{M}$  palmitate does not restore JNK responsiveness in LysM-FAS cells. **c–e**, Cholesterol content normalized by protein (**c**;  $n = 8$ ), cholesterol release to methyl- $\beta$ -cyclodextrin (**d**;  $n = 4$ ), and oxysterol synthesis in control and FAS-deficient cells (**e**;  $n = 5$ ). **f**, Control and LysM-FAS macrophages were loaded with  $0–5 \mu\text{M}$  cholesterol using methyl- $\beta$ -cyclodextrin followed by stimulation with LPS to assess levels of JNK/pJNK and RhoA. Representative blots are shown on the left and quantification of pJNK levels on the right ( $n = 5$ ). **g**, Western blot for pJNK in LysM-FAS macrophages after loading with  $1 \mu\text{M}$  cholesterol followed immediately by LPS stimulation (Chol.) or LPS stimulation after 2 h (Chol. Pulse). **h**, Confocal microscopy images of LysM-FAS cells expressing Rac-EGFP. Cells were subjected to vehicle or cholesterol loading followed by preparation of GPMVs and staining with FAST-Dil.

readily released to methyl- $\beta$ -cyclodextrin by LysM-FAS cells (Fig. 4d) and FAS-deficient RAW 264.7 cells (Extended Data Fig. 9g). LysM-FAS cells exhibited increased oxysterol synthesis in response to cholesterol loading (Fig. 4e), reflecting the increased metabolism of membrane cholesterol<sup>23</sup> and resulting in decreased cholesterol content. Intact FAS-deficient RAW 264.7 cells had increased accessibility to cholesterol oxidase (Extended Data Fig. 9h). Thus, endogenous fatty acid synthesis in macrophages regulates intracellular cholesterol metabolism.

Loading the cells with a range of cholesterol concentrations, using methyl- $\beta$ -cyclodextrin followed by exposure to LPS, revealed differences in JNK activation (Fig. 4f). As we expected, FAS-deficient cells exhibited decreased JNK activation when compared to control cells in the absence of cholesterol loading, although low concentrations of cholesterol rescued JNK activation. When the cholesterol pulse was followed by a two-hour cholesterol-free chase before addition of LPS, JNK activation was lost (Fig. 4g). Cholesterol loading of LysM-FAS cells increased Rac content in the liquid-ordered domain of GPMVs

(Fig. 4h–j) and loading the cells before preparing vesicles, or directly adding cholesterol to vesicles, corrected perturbations in membrane order (Fig. 4k). These results confirm that membrane order is cholesterol-dependent in GPMVs and is correlated with DRM cholesterol content<sup>24</sup>.

Loading macrophages with  $1 \mu\text{M}$  cholesterol rescued JNK activation by LPS in LysM-FAS macrophages, without significantly affecting JNK activation in controls (Fig. 4l). Loading with the enantiomer of cholesterol (*ent*-cholesterol<sup>25</sup>) rescued JNK activation with FAS deficiency (Fig. 4m), suggesting sterol interaction with membrane lipids rather than proteins. Loading with  $(3\beta)$ -26,27-dinorcholest-5-en-24-yn-3-ol, an alkyne with the ring structure of cholesterol but a modified side chain, also rescued JNK activation (Fig. 4n). However, loading with coprostanol, a non-planar sterol, did not (Fig. 4o). These results suggest that planar sterols (Extended Data Fig. 10a) interact with the FAS-deficient phospholipid environment to assemble signalling domains.

The endogenous synthesis of lipids in obesity-related diabetes relays physiological cues that cannot be mimicked by chemically identical exogenous lipids, consistent with studies of lipid metabolism in viral infection<sup>26</sup>. FAS configures the plasma membrane to retain cholesterol, which is required for propagating inflammatory signals. This may account for the beneficial effects of FAS inhibition in animal models of diabetes<sup>27</sup>. Without FAS (Extended Data Fig. 10b), intrinsic phospholipids incorporate unsaturated fatty acids that favour the release of cholesterol and Rho GTPases do not activate JNK. *De novo* lipogenesis may also configure the lipid environment to affect dendritic cell, T-helper 17 cell, and inflammasome function<sup>28–30</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 22 March; accepted 23 September 2016.**

**Published online 2 November 2016.**

- Olefsky, J. M. & Glass, C. K. Macrophages, inflammation, and insulin resistance. *Annu. Rev. Physiol.* **72**, 219–246 (2010).
- Ferrante, A. W. Jr. Macrophages, fat, and the emergence of immunometabolism. *J. Clin. Invest.* **123**, 4992–4993 (2013).
- Carvalho, J. B., Qiu, Y. & Chawla, A. Blood spotlight on leukocytes and obesity. *Blood* **122**, 3263–3267 (2013).
- Han, M. S. *et al.* JNK expression by macrophages promotes obesity-induced insulin resistance and inflammation. *Science* **339**, 218–222 (2013).
- Lumeng, C. N. & Saltiel, A. R. Inflammatory links between obesity and metabolic disease. *J. Clin. Invest.* **121**, 2111–2117 (2011).
- Köberlin, M. S. *et al.* A conserved circular network of coregulated lipids modulates innate immune responses. *Cell* **162**, 170–183 (2015).
- Foster, L. J., De Hoog, C. L. & Mann, M. Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc. Natl Acad. Sci. USA* **100**, 5813–5818 (2003).
- Yamashita, T. *et al.* Enhanced insulin sensitivity in mice lacking ganglioside GM3. *Proc. Natl Acad. Sci. USA* **100**, 3445–3449 (2003).
- Holzer, R. G. *et al.* Saturated fatty acids induce c-Src clustering within membrane subdomains, leading to JNK activation. *Cell* **147**, 173–184 (2011).
- Zhou, Y. *et al.* Membrane potential modulates plasma membrane phospholipid dynamics and K-Ras signaling. *Science* **349**, 873–876 (2015).
- Maier, T., Leibundgut, M. & Ban, N. The crystal structure of a mammalian fatty acid synthase. *Science* **321**, 1315–1322 (2008).
- Wei, X. *et al.* *De novo* lipogenesis maintains vascular homeostasis through endothelial nitric-oxide synthase (eNOS) palmitoylation. *J. Biol. Chem.* **286**, 2933–2945 (2011).
- Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R. & Mann, M. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat. Methods* **7**, 383–385 (2010).
- Simons, K. & Gerl, M. J. Revitalizing membrane rafts: new tools and insights. *Nat. Rev. Mol. Cell Biol.* **11**, 688–699 (2010).
- Gowrishankar, K. *et al.* Active remodeling of cortical actin regulates spatiotemporal organization of cell surface molecules. *Cell* **149**, 1353–1367 (2012).
- Lingwood, D. & Simons, K. Lipid rafts as a membrane-organizing principle. *Science* **327**, 46–50 (2010).
- Prag, S. *et al.* Activated ezrin promotes cell migration through recruitment of the GEF Dbl to lipid rafts and preferential downstream activation of Cdc42. *Mol. Biol. Cell* **18**, 2935–2948 (2007).
- Palazzo, A. F., Eng, C. H., Schlaepfer, D. D., Marcantonio, E. E. & Gundersen, G. G. Localized stabilization of microtubules by integrin- and FAK-facilitated Rho signaling. *Science* **303**, 836–839 (2004).
- Fessler, M. B. *et al.* Lipid rafts regulate lipopolysaccharide-induced activation of Cdc42 and inflammatory functions of the human neutrophil. *J. Biol. Chem.* **279**, 39989–39998 (2004).
- del Pozo, M. A. *et al.* Integrins regulate Rac targeting by internalization of membrane domains. *Science* **303**, 839–842 (2004).
- Sezgin, E. *et al.* Elucidating membrane structure and protein behavior using giant plasma membrane vesicles. *Nat. Protocols* **7**, 1042–1051 (2012).
- Levental, I., Grzybek, M. & Simons, K. Raft domains of variable properties and compositions in plasma membrane vesicles. *Proc. Natl Acad. Sci. USA* **108**, 11411–11416 (2011).
- Lange, Y. *et al.* Regulation of fibroblast mitochondrial 27-hydroxycholesterol production by active plasma membrane cholesterol. *J. Lipid Res.* **50**, 1881–1888 (2009).
- Levental, I. *et al.* Cholesterol-dependent phase separation in cell-derived giant plasma-membrane vesicles. *Biochem. J.* **424**, 163–167 (2009).
- Westover, E. J. & Covey, D. F. The enantiomer of cholesterol. *J. Membr. Biol.* **202**, 61–72 (2004).
- York, A. G. *et al.* Limiting cholesterol biosynthetic flux spontaneously engages type I IFN signaling. *Cell* **163**, 1716–1729 (2015).
- Wu, M. *et al.* Antidiabetic and antisteatotic effects of the selective fatty acid synthase (FAS) inhibitor platensimycin in mouse models of diabetes. *Proc. Natl Acad. Sci. USA* **108**, 5378–5383 (2011).
- Everts, B. *et al.* TLR-driven early glycolytic reprogramming via the kinases TBK1-IKKε supports the anabolic demands of dendritic cell activation. *Nat. Immunol.* **15**, 323–332 (2014).
- Berod, L. *et al.* *De novo* fatty acid synthesis controls the fate between regulatory T and T helper 17 cells. *Nat. Med.* **20**, 1327–1333 (2014).
- Moon, J. S. *et al.* UCP2-induced fatty acid synthase promotes NLRP3 inflammasome activation during sepsis. *J. Clin. Invest.* **125**, 665–680 (2015).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by NIH grants DK101392, DK076729, DK088083, DK20579, DK56341, RR00954, HL067773 and the Taylor Family Institute for Innovative Psychiatric Research. S. Teitelbaum provided the constitutively active Rac construct, T. Pryse performed confocal microscopy, L. Yang and M. Miller assisted with imaging and L. Mydock-McGrane prepared the cholesterol alkyne.

**Author Contributions** X.W., H.S., L.Y., M.G.R., and R.S. performed the experiments. H.S. performed lipidomic and proteomic analyses. R.S. and D.S.O. designed and performed the analysis for oxysterols. D.F.C. designed experiments using sterol analogues. X.W. and C.F.S. designed the experiments, analysed all data, and wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.F.S. ([csemenko@wustl.edu](mailto:csemenko@wustl.edu)).

**Reviewer Information** Nature thanks A. Chawla, A. Tall and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

**Animal models.** Haematopoietic FAS ablation mouse models LysM-FAS<sup>31</sup>, Tie2-FAS<sup>32</sup>, and Tie2-FAS bone-marrow transplantation animals<sup>12</sup> were generated as described previously. The high-fat diet was a Western-type diet containing 0.15% cholesterol with 42% calories as fat (TD 88137, Harlan). Studies were conducted with littermates of the C57BL/6 background between 2 and 6 months of age in a specific pathogen-free facility with a 12-h light–12-h dark cycle. The Animal Studies Committee at Washington University in St. Louis approved experiments.

Sample sizes were based on variance of previous studies in similar experimental settings. A formal randomization tool was not used to allocate animals to experimental groups. Cages containing both genotypes as littermates were selected for allocation to experimental groups. Animals were not excluded from analyses unless results could not be generated due to the death of the animal. No formal blinding was employed in the animal studies. For metabolic phenotyping experiments, males were studied. For cell biology experiments, macrophages from both genders were studied.

**Metabolic phenotyping.** For glucose tolerance tests, mice were injected intraperitoneally (i.p.) with 1 g kg<sup>-1</sup> glucose after 6 h of fasting; glucose was measured in tail blood using a Contour glucometer (Bayer). Glucose-stimulated insulin-secretion assays were performed in separate cohorts, with insulin measured using ELISA kits (PerkinElmer). For insulin tolerance tests, mice were injected i.p. with 0.75 U kg<sup>-1</sup> insulin after 6 h of fasting. For insulin signalling, mice were injected i.p. with 5 mU g<sup>-1</sup> insulin after overnight fasting and tissues were collected 10 min later. Body composition analysis was performed using magnetic resonance imaging. Liver triglycerides<sup>32</sup> and food intake<sup>12</sup> were assayed as described previously. Global insulin sensitivity was determined by hyperinsulinaemic–euglycaemic clamp as described previously<sup>33</sup>. Animals were implanted with a jugular catheter. Five days after surgery, animals were fasted for 4 h and glucose turnover was measured in the basal state and during the clamp in conscious mice.

**Histology.** Visceral fat was fixed in formalin. Sections of 5 µm were stained for Mac2 (BD Bioscience) for crown-like structures. Processing included slides not treated with primary antibodies to correct for any non-specific staining. For neutral lipid detection, 10-µm sections of frozen liver were fixed with formalin (10%, Sigma) followed by Oil red O staining (Sigma) in 60% isopropanol.

**Isolation of stromal-vascular cell fraction (SVF) and FACS analysis.** Epididymal fat pads were isolated after saline perfusion, minced, washed, and centrifuged (1,000 g for 5 min). Floating adipose tissue was collected and digested with type I collagenase solution (1 mg ml<sup>-1</sup>, 30 min). This digest was pelleted to yield the SVF, which was washed twice in FACS wash buffer (PBS supplemented with 4% FBS), then blocked for Fc receptors using anti-mouse CD16/32 (BD Bioscience), followed by staining with APC-conjugated anti-mouse CD11b, eFluor 450-conjugated anti-mouse CD11c, and PE-conjugated anti-mouse F4/80 (eBioscience). Other surface markers included CD18 and ICAM-1 (both PE-conjugated, eBioscience). Samples were analysed using a BD LSII flow cytometer.

**Gene expression.** Total RNA was extracted using TRIzol (Invitrogen) and reverse transcribed using an iScript cDNA synthesis kit (Bio-Rad). PCR reactions were performed with an ABI PRISM 7000 Sequence Detection System (Applied Biosciences) using the SYBR Green PCR Master Mix assay and primer sequences as described previously<sup>34</sup>.

**Western blotting and Rho GTPase activity assays.** Tissues or cultured cells were treated with lysis buffer (50 mM Tris HCl, pH 7.4, 1 mM EDTA, 150 mM NaCl, 1% NP40, 0.25% Na deoxycholate, 2 mM NaVO<sub>3</sub>, 5 mM NaF, and protease inhibitors; Roche). Whole lysates or fractions were subjected to 4–20% gradient SDS–PAGE (Invitrogen) followed by immunoblotting with antibodies against FAS, transferrin receptor (Abcam), actin (Sigma), total and pAkt (residues T308 and S473), total and pJNK, Lyn, RhoA, Cdc42, Rac, moesin (Cell Signaling), flotillin-1, insulin receptor (BD Bioscience), ICAM-1, myosin (Santa Cruz), and annexin V (Biosensis). Activity of Rho family GTPases was measured by pull-down assays for GTP-bound forms of the proteins (Cytoskeleton). Lysates were mixed with PAK–GST protein beads (for GTP–Rac binding) or Rhotekin–RBD protein GST beads (for GTP–RhoA binding) and precipitates were analysed by immunoblotting.

**Cell culture and metabolic labelling.** Bone-marrow-derived macrophages were differentiated in DMEM with 20% L929-conditioned medium. Peritoneal macrophages were elicited from mice by i.p. injection of 4% thioglycollate medium (Sigma), and adherent cells were cultured in DMEM with 10% FBS. RAW 264.7 cells and 293T cells were obtained from the ATCC. RAW 264.7 cells were cultured in DMEM plus 10% FBS. FAS was knocked down in RAW 264.7 cells using a lentiviral-based shRNA strategy (Open Biosystems). 293T cells were transfected with packaging vectors, along with an expression plasmid (pLKO.1-puro system) containing shRNA sequences that were scrambled or specific for mouse FAS mRNA<sup>34</sup>. Viruses were collected and filtered two days later, then used to infect

RAW 264.7 cells (10 µg ml<sup>-1</sup> polybrene). Infected RAW 264.7 cells were selected with puromycin (4 µg ml<sup>-1</sup>) for 2 days.

For inflammatory activation, macrophages were treated with LPS (100 ng ml<sup>-1</sup>), or high-dose palmitate (500 µM/1% BSA, with BSA only as control) for 6 h before collection. Lysates were assayed for JNK activation by western blotting. Supernatants were assayed for the pro-inflammatory cytokines TNFα, MCP1, IL1β, and IL12p40 by ELISA (R&D systems).

For lipid labelling analyses, cells were serum-starved for 24 h in the absence or presence of simvastatin (10 µM of the active form, Calbiochem). <sup>14</sup>C-acetate (10 µCi) was added to 10-cm culture dishes, cells were collected after 4 h, lipids were extracted with chloroform/methanol and radioactivity was measured with a liquid scintillation counter. In chase experiments involving *de novo* lipid synthesis, cells were treated with <sup>14</sup>C-acetate (10 µCi) for 24 h then chased with cold medium for 0.5, 1 or 4 h. DRMs were isolated and lipid extracts were counted. In chase experiments involving palmitate, cells were incubated with <sup>14</sup>C-palmitate (5 µCi, ~10 µM) complexed with cold palmitate (50 µM)/BSA (0.1%) then chased with cold medium. In exogenous palmitate reconstitution experiments, FAS-knockout macrophages were treated with palmitate (50 µM/0.1% BSA) for 24 h with BSA-only groups serving as controls. Beyond determination of FAS status, cell identity was not authenticated and cells were not tested for mycoplasma.

**Membrane extraction and isolation of submembrane microdomains.** Cells were homogenized in hypotonic buffer (1 mM HEPES, pH 8.0, 15 mM KCl, 2 mM MgCl<sub>2</sub>, 0.1 mM EDTA, and protease inhibitors from Roche). Lysates were subjected to sequential centrifugation steps (2,000 g for 5 min, 10,000 g for 15 min, 100,000 g for 2 h) to yield crude membrane fractions. To isolate detergent-resistant membranes (DRMs), cells were lysed in MES buffer (10 mM MES, 150 mM NaCl, pH 6.5) containing protease inhibitors, incubated with 1% Triton X-100 on ice for 30 min, and homogenized by 20 passes through 29-gauge needles. Homogenates were adjusted to 40% sucrose and placed under sucrose layers of 5% and 30%. After centrifugation at 39,000 r.p.m., 4°C for 16 h, fractions were collected from top to bottom. For submembrane isolation in the absence of detergent, sodium carbonate buffer at high pH (500 mM, pH 11.0) was used and Triton X-100 was omitted in the above procedure. Sucrose layers of 5%, 35% and 45% were used to generate fractions.

**Lipid analyses by electrospray ionization–tandem mass spectrometry.** After Bligh–Dyer extraction, organic phases were collected, dried under nitrogen, and reconstituted in 200 µl chloroform/methanol (1:1) with 0.5% sodium acetate. A 50-µl aliquot was directly injected into a Thermo Vantage triple–quadrupole mass spectrometer in positive mode for the analysis of phosphatidylcholine (including sphingomyelin) species with neutral loss scan of 183, and for the analysis of phosphatidylethanolamine species with neutral loss scan of 141. Each individual species was compared to its internal standard, and absolute quantity was determined using a standard curve, all as described<sup>33</sup>.

**SILAC and proteomic analysis by mass spectrometry.** SILAC techniques<sup>35,36</sup> used RAW 264.7 cells cultured in medium containing heavy (that is, <sup>13</sup>C<sub>6</sub> L-lysine and <sup>13</sup>C<sub>6</sub> L-arginine) stable-isotope-labelled amino acids (Thermo Scientific). Murine bone-marrow-derived macrophage cells from different genetic models were grown in light (that is, <sup>12</sup>C<sub>6</sub> L-lysine and <sup>12</sup>C<sub>6</sub> L-arginine) medium. Cells were lysed and equal amounts of extracts from labelled cells were combined, then subjected to differential centrifugation for the crude membrane extraction, or sucrose gradient centrifugation for DRM extraction. Samples were resolved by SDS–PAGE and separated into 10 fractions, and in-gel trypsin digestion was performed before liquid chromatography–tandem mass spectrometry (LC–MS/MS).

For nano-high-pressure liquid chromatography–electrospray ionization–MS/MS, studies were performed on a LTQ Orbitrap (Thermo) instrument. Samples were loaded with an autosampler onto a 15-cm Magic C18 column (5-µm particles, 300-Å pores, Michrom Bioresources) packed into a PicoFrit tip (New Objective) and analysed with 2D nanoLC plus HPLC (Eksigent). Analytical gradients were from 0–80% organic phase (95% acetonitrile, 0.1% formic acid) over 60 min. Aqueous phase composition was 2% acetonitrile, 0.1% formic acid. Eluent was routed into a PV-550 Nanospray ion source (New Objective). The LTQ Orbitrap instrument was operated in a data-dependent mode with the precursor scan over the range *m/z*: 350–2,000, followed by twenty MS2 scans using parent ions selected from the MS1 scan. The Orbitrap AGC target was set to 1 × 10<sup>6</sup>, and the MS2 AGC target was 1 × 10<sup>4</sup> with maximum injection times of 300 ms and 500 ms, respectively. For MS/MS, LTQ isolation width was 2 Da, normalized collision energy was 30% and activation time was 10 ms.

Raw data were submitted to Mascot Server 2.0 and searched against the SwissProt database. Results were quantified by analysing the mascot ‘dat’ file and its respective thermo ‘raw’ file using a locally generated program. Relative protein ratios for control versus knockout macrophages were calculated for each identified protein (averaging the signal from multiple peptides) and presented as a percentage



of the control in a heat map (CIMminer). Pathway analysis was performed using databases including DAVID, KEGG and PANTHER.

**Analysis of GPMVs.** GPMVs were induced in bone-marrow-derived macrophages or RAW 264.7 cells essentially as described previously<sup>21</sup>. Cells were rinsed with GPMV buffer (10 mM HEPES, 150 mM NaCl and 2 mM CaCl<sub>2</sub>, pH 7.4) followed by a 1–2-h incubation with vesiculation-induction GPMV buffer in the presence of PFA (25 mM) and DTT (2 mM) at 37°C. The vesicles were stained using FAST-DiI (Invitrogen, 0.25 µg ml<sup>-1</sup>) at room temperature for 30 min. Vesicles were imaged during temperature-controlled cooling by confocal microscopy<sup>37</sup>. A four-chambered cover glass containing vesicle suspensions was placed in a metal block that was gradually cooled by circulating water. Images were captured with a 20× air objective in an inverted Zeiss LSM 510 laser-scanning microscope system. An HeNe laser (at 543 nm excitation) was used for FAST-DiI. In multiple scanning mode for sequential captures, an Argon laser (at 488 nm excitation) was used for EGFP imaging, and single-channel labelled cells were used to calibrate the system to exclude the leak between the two channels.

For lipid order analysis, an environmentally sensitive dye di-4-ANEPPDHQ was used as described<sup>38</sup>. The dye was excited by an Argon laser at 488 nm wavelength and the two-channel emission signal was collected at ~560 nm (liquid ordered) and ~620 nm (liquid disordered) simultaneously. Image calculations were carried out in ImageJ using the GP (generalized polarization analysis) plugin ([http://www.optinav.com/Generalized\\_Polarization\\_Analysis.htm](http://www.optinav.com/Generalized_Polarization_Analysis.htm)) with modification to include a measured *G* (calibration factor). To determine the general polarization (GP) value, we analysed the pictures acquired and calculated the general polarizations using the equation:

$$GP = \frac{I_{blue} - G I_{red}}{I_{blue} + G I_{red}}$$

where *G* is the calibration factor calculated using

$$G = \frac{GP_{ref} + GP_{ref}GP_{mes} - GP_{mes} - 1}{GP_{mes} + GP_{ref}GP_{mes} - GP_{ref} - 1}$$

*GP<sub>mes</sub>* is the general polarization value of the dye in 100% DMSO measured using the same imaging settings as used for GPMVs.

To analyse the partition patterns of Rac protein in GPMVs, cells expressing Rac-EGFP were used. A retroviral-based plasmid for EGFP conjugated wild-type Rac was derived from the pMX-GFP-Rac<sup>G12V</sup> (Addgene #14567, N-terminal GFP) by site-directed mutagenesis. The plasmid was transfected into PLAT-E cells, and a retrovirus vector was used to infect primary bone marrow progenitor cells during macrophage differentiation. The GPMVs generated were then labelled with FAST-DiI for monitoring phase separation, and two-channel fluorescence was captured at 15°C. The GFP fluorescence intensity peaks in the two domains were quantified. Analysis of the temperature plot of phase separation and quantification of Rac protein membrane partition was performed using ImageJ software<sup>37</sup>.

**Cell spreading and migration assays.** For spreading, macrophages were plated (30,000 cells per well) into four-chambered Lab-Tek slides (Thermo Scientific). After 30 min, cells were gently washed and fixed with 4% paraformaldehyde, followed by permeabilization with 0.1% Triton X-100. The cells were then stained with rhodamine-phalloidin (Invitrogen). Cell images were captured by immunofluorescence microscopy. Individual cells were outlined, and total cell area was quantified using ImageJ software.

For migration, Transwell inserts with 3-µm pore size (Corning) were pre-coated with 0.2% gelatin. Macrophages were trypsinized and 20,000 cells were added in triplicate to inserts in chambers. Medium with vehicle or MCP1 (100 ng ml<sup>-1</sup>, R&D systems) was added to the lower wells and, 4 h later, cells that had migrated to the underside of the membrane were fixed. Cells on the upper side of the membrane were removed, and membranes were cut and positioned with migrated cells facing up, followed by DAPI staining and counting.

**Sterol manipulations.** Cholesterol oxidase activity and membrane cholesterol release potential were measured as described previously<sup>39</sup>. Cells cultured overnight in 96-well plates were rinsed twice with PBS then treated with 100 µl of PBS containing cholesterol oxidase (2 U ml<sup>-1</sup>, Sigma) at 37°C for 10 min. Then 50 µl Amplex red reagent (Invitrogen) was added, and after incubation at 37°C for 20 min, activity was quantitated by spectrometry at 560 nm. Samples processed without cholesterol oxidase were used to determine background.

For cholesterol release potential, cells were labelled with <sup>3</sup>H-cholesterol overnight, rinsed with cold medium three times then incubated with cold medium containing methyl-β-cyclodextrin (1 mM, Sigma) at 37°C. Aliquots of the medium were collected over time, sedimented and the supernatants were counted. Cell lysates were also processed and counted.

For oxysterol production with cholesterol loading<sup>23</sup>, cells were plated in 6-well plates (50,000 cells/well) and incubated overnight in medium with

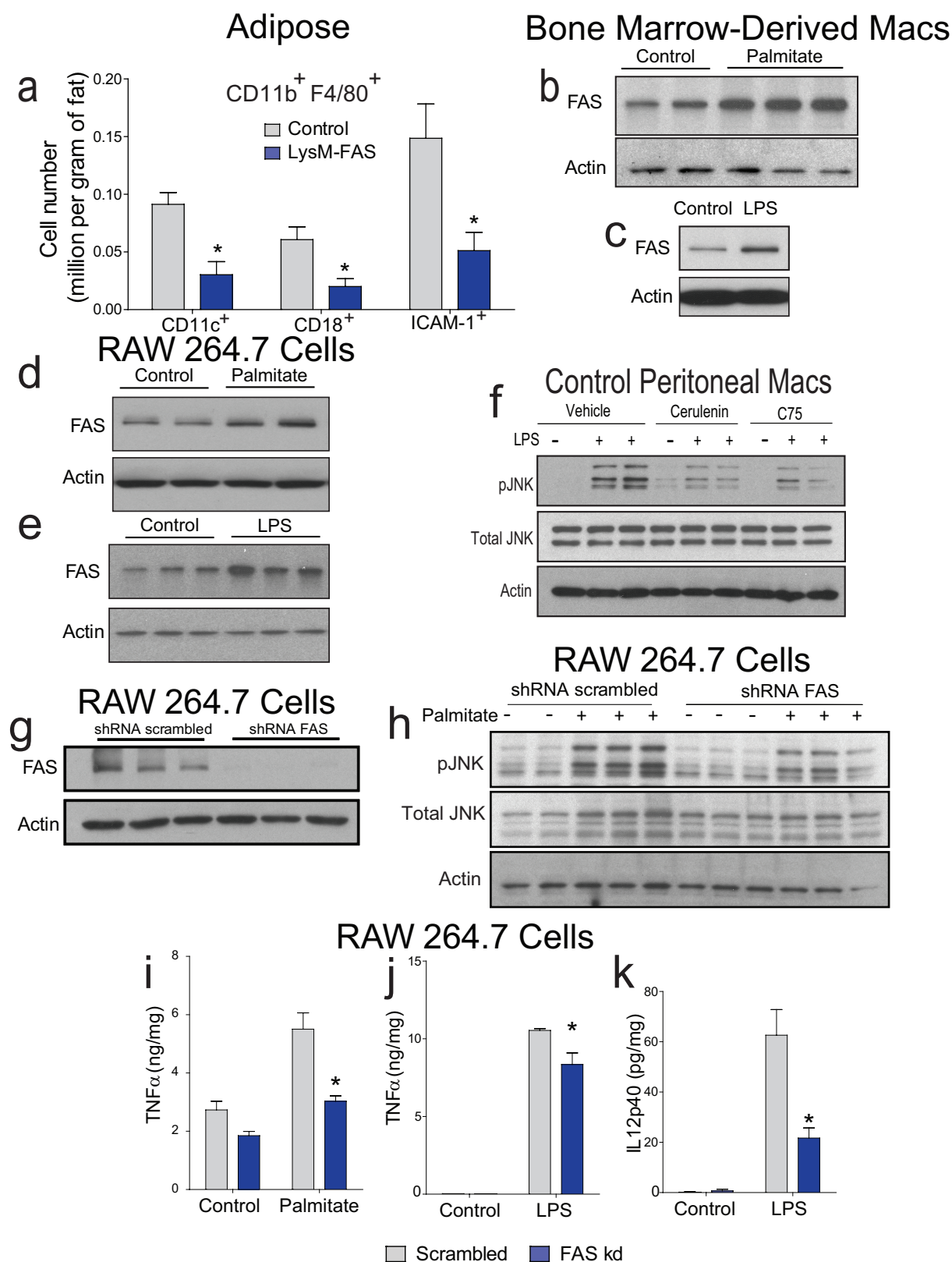
lipoprotein-deficient serum. Cells were treated with a methyl-β-cyclodextrin: cholesterol complex for 10 min, then washed twice and harvested for oxysterol measurements<sup>40</sup>. 25-hydroxycholesterol (25-HC) and 27-hydroxycholesterol (27-HC) were extracted by the Bligh–Dyer method from homogenized macrophages and medium after addition of deuterated internal standard (d5-27-HC). The organic layer was taken to dryness under nitrogen, then 50 µl of 0.5 M *N,N*-dimethylglycine/2M 4-dimethylaminopyridine in chloroform and 50 µl of 1 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide in chloroform were added to derivatize the samples. After incubation for 1 h at 50°C, the reaction was quenched with 50 µl of methanol. The sample was taken to dryness under nitrogen and the residue was treated with 1:3 (v/v) water:hexane to remove the derivatizing reagent. The hexane layer was taken to dryness under nitrogen and the sample was reconstituted with 200 µl of methanol and analysed by LC–MS/MS using a Prominence HPLC system (Shimadzu Scientific Instruments) and a 4000QTRAP mass spectrometer (Applied Biosystems/MDS Sciex Inc.). Data were acquired using Analyst software (v.1.5.1).

**Rescue of JNK signalling.** To express a constitutively active form of Rac protein in FASc-knockout macrophages, we used a selectable retroviral expression system using pMX-Rac<sup>G12V</sup> with a blasticidin-resistance gene. Plasmids were transfected into PLAT-E cells, and produced retrovirus was used to infect primary bone marrow progenitor cells during macrophage differentiation. On the day following infection, positive cells were selected in blasticidin (1 µg ml<sup>-1</sup>). Differentiated macrophages were used for analysing JNK response 3 days later.

For cholesterol rescue experiments, different forms and concentrations of cholesterol were complexed to methyl-β-cyclodextrin (Sigma) at a ratio of 1:10 (with 0.25 mM cyclodextrin) and diluted to desired concentrations using serum-free medium. *ent*-cholesterol<sup>25</sup> and alkyne cholesterol<sup>41</sup> were synthesized at Washington University. Coprostanol was purchased from Sigma. Sterol was first dried under nitrogen gas, and then sonicated into DMEM with cyclodextrin until flaky chunks disappeared. The solution was then shaken overnight at 37°C. Bone-marrow-derived macrophages were pretreated with serum-free medium for 2 h, and then cyclodextrin/cholesterol was added for 10 min followed by 30 min exposure to LPS (100 ng ml<sup>-1</sup>). For the 'cholesterol pulse' experiment, a two-hour serum-free medium incubation followed cyclodextrin/cholesterol exposure before LPS was added. For reconstitution of cholesterol in GPMVs, cells were treated with cholesterol (25 µM) overnight followed by vesicle preparation. For some experiments, cholesterol solution was added directly to vesicle suspensions at 5 µg ml<sup>-1</sup>. **Statistical analyses.** Data are expressed as mean ± s.e.m. Analyses were performed with GraphPad Prism by two-tailed *t*-test (for two groups), one-way ANOVA (more than two groups) and Tukey's multiple comparison test, by two-way ANOVA (two independent variables) and Bonferroni post-tests, or by nonlinear curve-fit comparison (where indicated). *P* < 0.05 was considered significant and is indicated with an asterisk (except for results of nonlinear curve fit comparisons, where the asterisk indicates *P* < 0.001). For analysis of proteomic data, pathways with *P* < 0.05 (modified Fisher's exact *P* value for gene-enrichment analysis) were selected for protein–protein interaction maps generated by STRING.

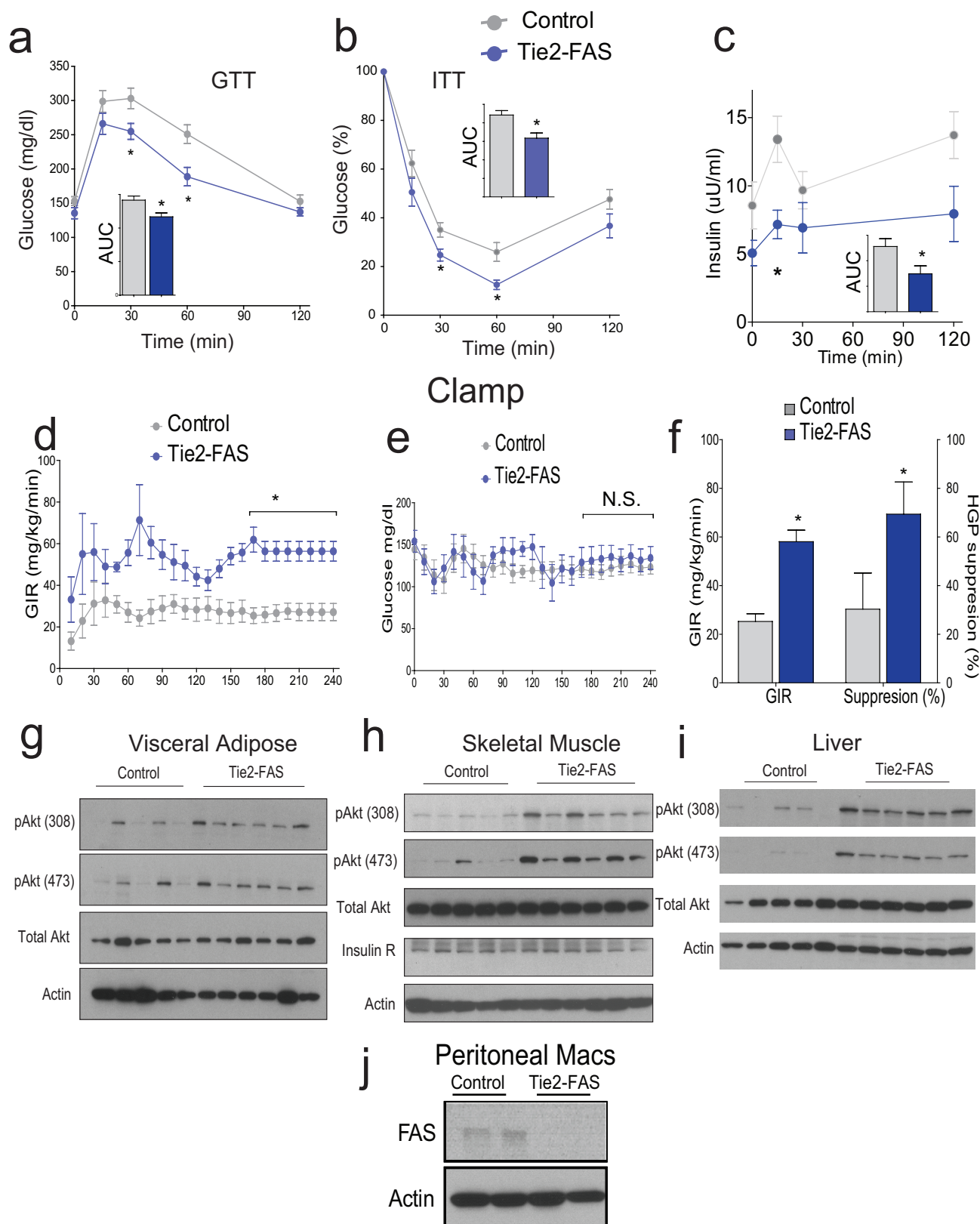
**Data availability.** Data that support the findings of this study are available from the corresponding author upon reasonable request.

- Schneider, J. G. *et al.* Macrophage fatty-acid synthase deficiency decreases diet-induced atherosclerosis. *J. Biol. Chem.* **285**, 23398–23409 (2010).
- Chakravarthy, M. V. *et al.* Identification of a physiologically relevant endogenous ligand for PPARα in liver. *Cell* **138**, 476–488 (2009).
- Funai, K. *et al.* Muscle lipogenesis balances insulin sensitivity and strength through calcium signaling. *J. Clin. Invest.* **123**, 1229–1240 (2013).
- Wei, X. *et al.* Fatty acid synthase modulates intestinal barrier function through palmitoylation of mucin 2. *Cell Host Microbe* **11**, 140–152 (2012).
- Monetti, M., Nagaraj, N., Sharma, K. & Mann, M. Large-scale phosphosite quantification in tissues by a spike-in SILAC method. *Nat. Methods* **8**, 655–658 (2011).
- Wei, X., Song, H. & Semenkovich, C. F. Insulin-regulated protein palmitoylation impacts endothelial cell function. *Arterioscler. Thromb. Vasc. Biol.* **34**, 346–354 (2014).
- Levental, K. R. & Levental, I. Isolation of giant plasma membrane vesicles for evaluation of plasma membrane structure and protein partitioning. *Methods Mol. Biol.* **1232**, 65–77 (2015).
- Owen, D. M., Rentero, C., Magenau, A., Abu-Siniyeh, A. & Gaus, K. Quantitative imaging of membrane lipid order in cells and organisms. *Nat. Protocols* **7**, 24–35 (2011).
- Lange, Y., Ye, J. & Steck, T. L. How cholesterol homeostasis is regulated by plasma membrane cholesterol in excess of phospholipids. *Proc. Natl Acad. Sci. USA* **101**, 11664–11667 (2004).
- Jiang, X. *et al.* A sensitive and specific LC–MS/MS method for rapid diagnosis of Niemann-Pick C1 disease from human plasma. *J. Lipid Res.* **52**, 1435–1445 (2011).
- Cieplak, P. *et al.* New chemical probes targeting cholesterylolation of Sonic Hedgehog in human cells and zebrafish. *Chem. Sci.* **5**, 4249–4259 (2014).



**Extended Data Figure 1 | FAS is induced by inflammatory stimuli and its deficiency suppresses inflammation.** **a**, Quantification of subsets of pro-inflammatory macrophages (additionally positive for CD11c, CD18, or ICAM1) in the adipose tissue from HFD-fed control and LysM-FAS mice ( $n = 4$ ). **b**, **c**, Western blotting of FAS protein in bone-marrow-derived macrophages treated with vehicle, palmitate (**b**; 500  $\mu$ M) or LPS (**c**; 100  $\text{ng ml}^{-1}$ ). **d**, **e**, Western blot of FAS protein in RAW 264.7 cells treated with vehicle as control, palmitate (**d**; 500  $\mu$ M) or LPS (**e**; 100  $\text{ng ml}^{-1}$ ).

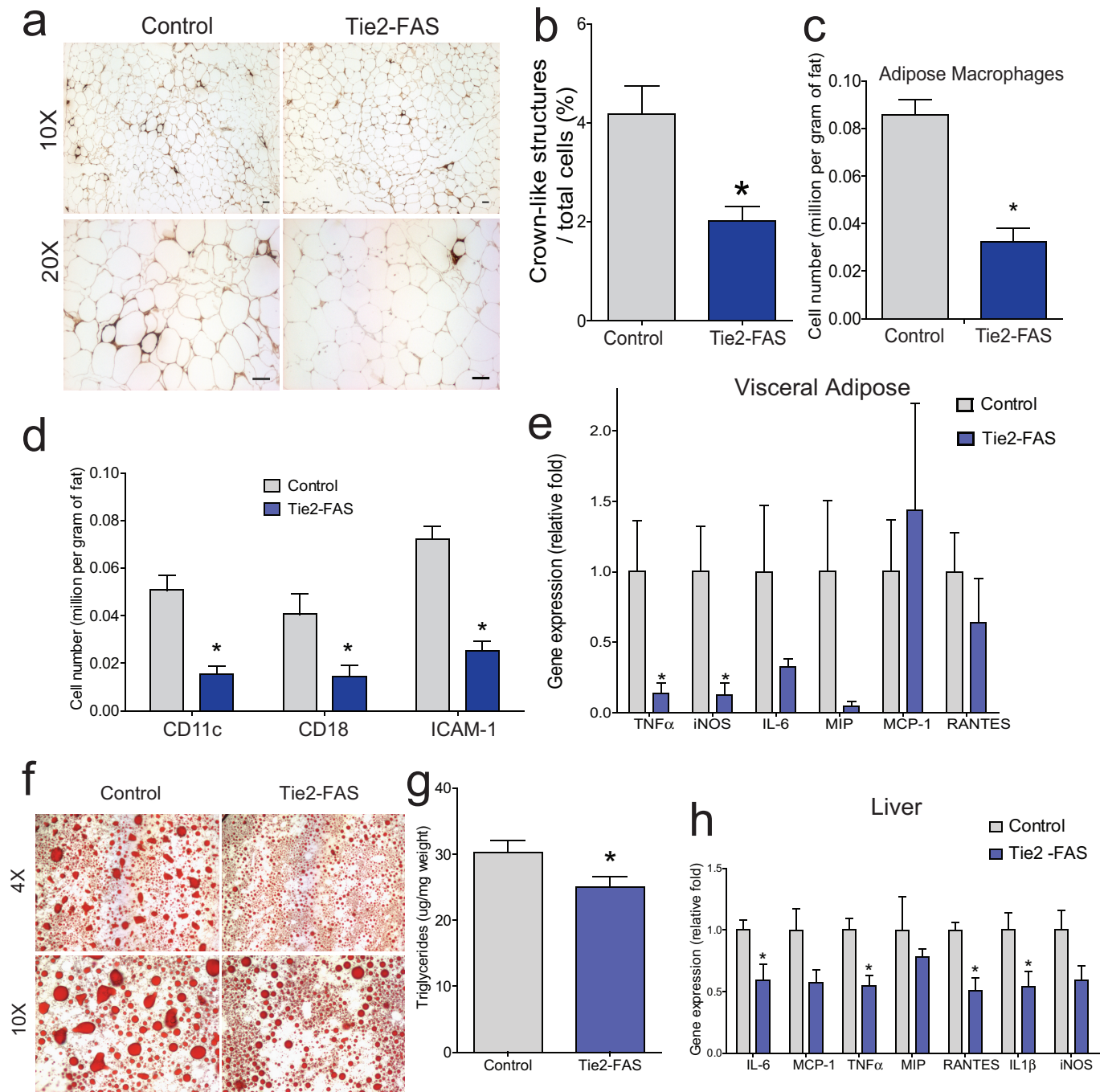
**f**, JNK phosphorylation levels after LPS stimulation in the presence of the FAS inhibitors cerulein or C75. **g**, Western blotting of FAS protein in RAW 264.7 cells treated with scrambled control or a lentiviral-based shRNA targeting FAS. **h**, Western blot of pJNK in RAW 264.7 cells after stimulation with palmitate (500  $\mu$ M). **i**–**k**, ELISAs for pro-inflammatory cytokines in RAW 264.7 cells stimulated with palmitate (**i**; 500  $\mu$ M) or LPS (**j**, **k**; 100  $\text{ng ml}^{-1}$ ),  $n = 3$ . \* $P < 0.05$ , data are mean  $\pm$  s.e.m.



**Extended Data Figure 2 | Tie2-FAS mice are protected from diet-induced insulin resistance.** **a–c**, Glucose tolerance test (**a**), insulin tolerance test (**b**) and glucose-stimulated insulin secretion assays (**c**) in control and Tie2-FAS mice fed with HFD for 3 months ( $n = 15$  control mice and 8 Tie2-FAS). **d**, **e**, Glucose infusion rate (GIR) (**d**), and plasma glucose levels (**e**) during hyperinsulinaemic–euglycaemic clamping for

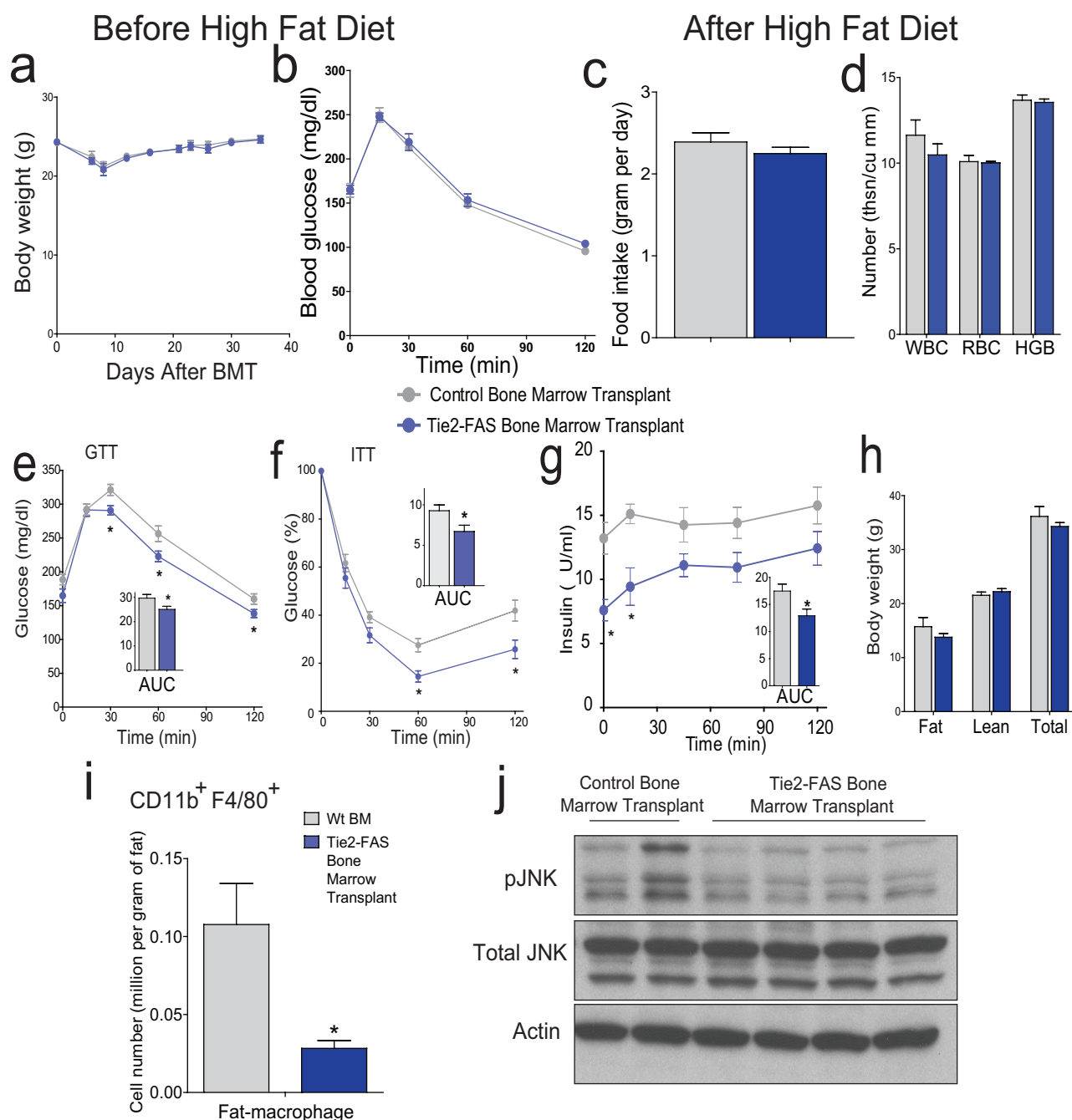
HFD-fed control ( $n = 9$ ) and Tie2-FAS mice ( $n = 6$ ). **f**, Mean GIR and hepatic glucose production (HGP) suppression in the steady state of the clamp experiment. **g–i**, Western blotting of Akt phosphorylation levels in various indicated mouse tissues during the clamp. **j**, Western blotting of FAS protein in peritoneal macrophages from control and Tie2-FAS mice. \* $P < 0.05$ , data are mean  $\pm$  s.e.m.





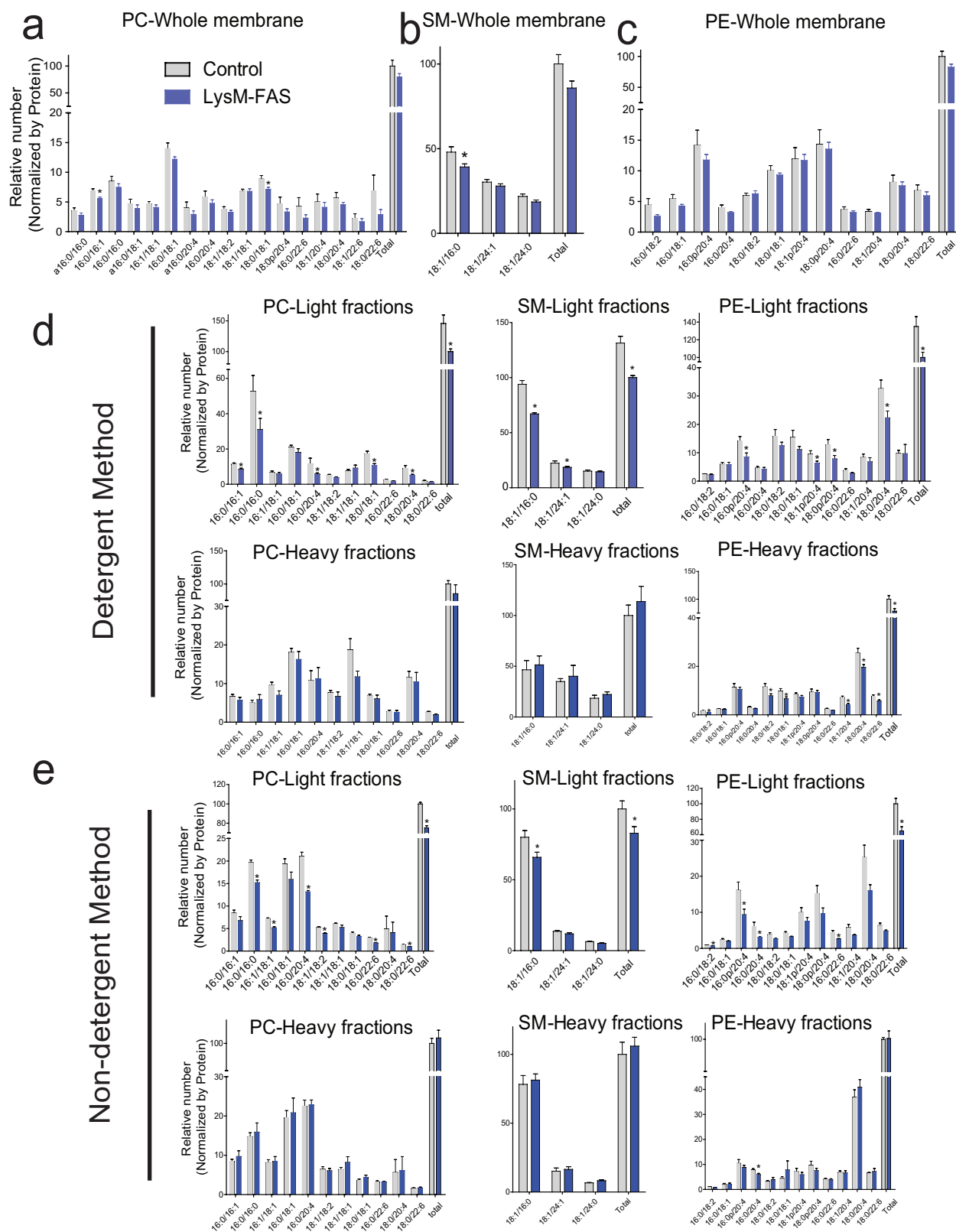
**Extended Data Figure 3 | Mice with FAS deficiency induced by Tie2-Cre resist diet-induced chronic inflammation.** **a**, Images of crown-like structures in visceral fat stained for the macrophage marker Mac2. **b**, Quantification of crown-like structures ( $n = 4$  mice). **c**, **d**, Quantification of macrophages (F4/80 $^{+}$  CD11b $^{+}$ ; **c**) and subsets of pro-inflammatory macrophages (those that are additionally positive for CD11c, CD18 and ICAM1; **d**) in visceral adipose tissue of Tie2-FAS mice by flow cytometry

( $n = 4$  control and 5 Tie2-FAS mice). **e**, Gene expression in visceral adipose tissue from control ( $n = 6$ ) and Tie2-FAS mice ( $n = 9$ ). **f**, Oil red O staining of liver from Tie2-FAS mice fed HFD. **g**, Quantification of liver fat content ( $n = 8$  control and 9 Tie2-FAS mice). **h**, Gene expression in liver from control ( $n = 6$ ) and Tie2-FAS mice ( $n = 9$ ). \* $P < 0.05$ , data are mean  $\pm$  s.e.m.



**Extended Data Figure 4 | Diet effects in Tie2-FAS bone marrow transplant mice.** **a**, Body weight measurements in mice ( $n=15$ ) following bone marrow transplant (BMT) with control or Tie2-FAS marrow on a chow diet. **b**, Glucose tolerance testing in mice ( $n=10$ ) one month after bone marrow transplant on a chow diet. **c**, **d**, Food intake (**c**) and blood counts (**d**) in mice on HFD following bone marrow transplant ( $n=7$  control mice and 9 Tie2-FAS bone marrow transplanted). **e**–**h**, Glucose

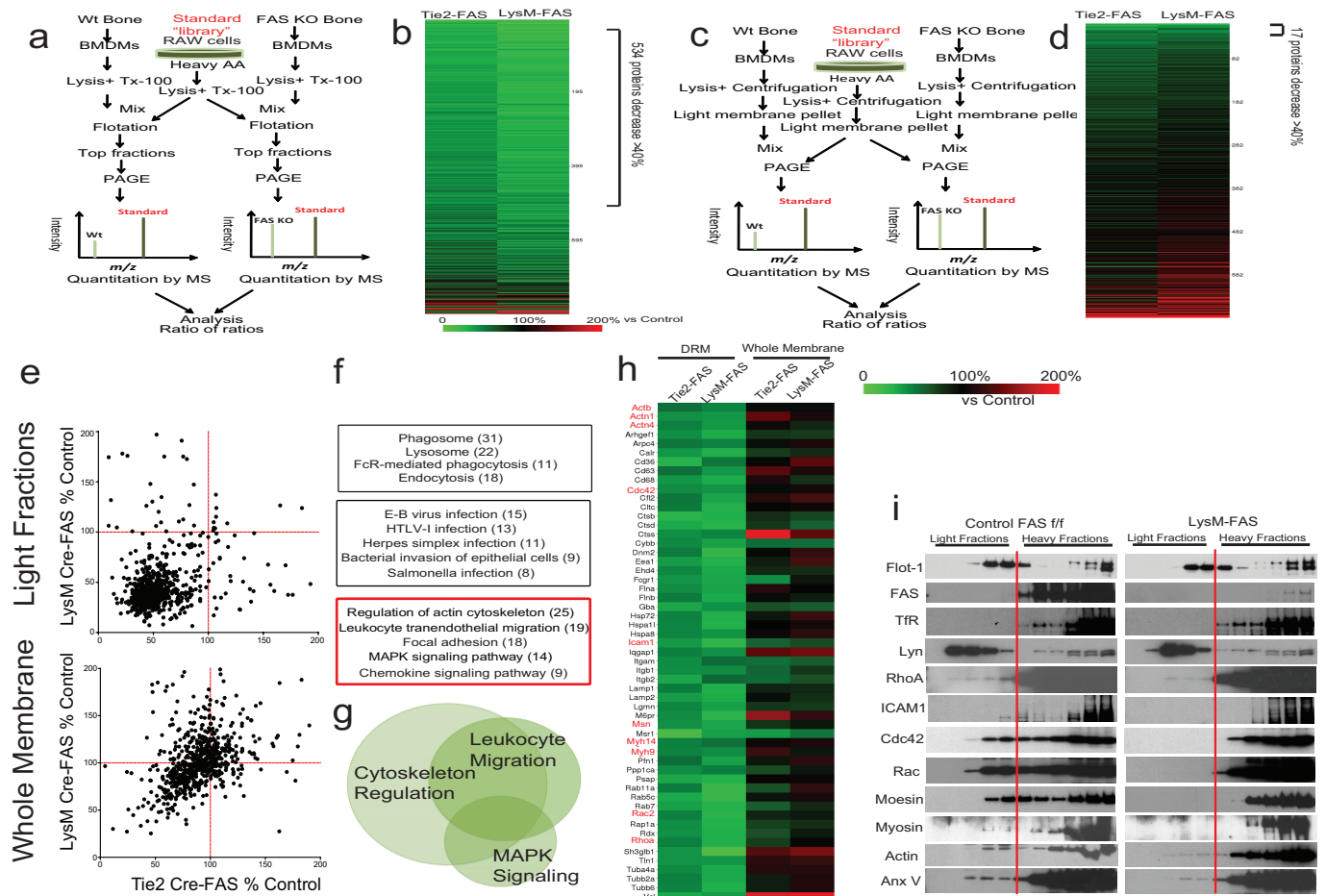
tolerance testing (**e**), insulin tolerance testing (**f**), glucose-stimulated insulin levels (**g**) and body composition (**h**) in control ( $n=17$ ; grey) and Tie2-FAS bone marrow transplant mice ( $n=15$ ; blue) on a HFD. **i**, Number of inflammatory cells in visceral adipose tissue ( $n=4$ ). **j**, JNK phosphorylation in adipose tissue of bone marrow transplant mice. \* $P < 0.05$ , data are mean  $\pm$  s.e.m.



**Extended Data Figure 5 | Phospholipid composition of light fractions (DRMs) and heavy fractions (non-DRM) in the presence and absence of FAS in macrophages.** **a–c**, Mass spectroscopic quantification of the relative number of phosphatidylcholine (**a**; PC), sphingomyelin (**b**; SM) and phosphatidylethanolamine (**c**; PE) species of phospholipids in

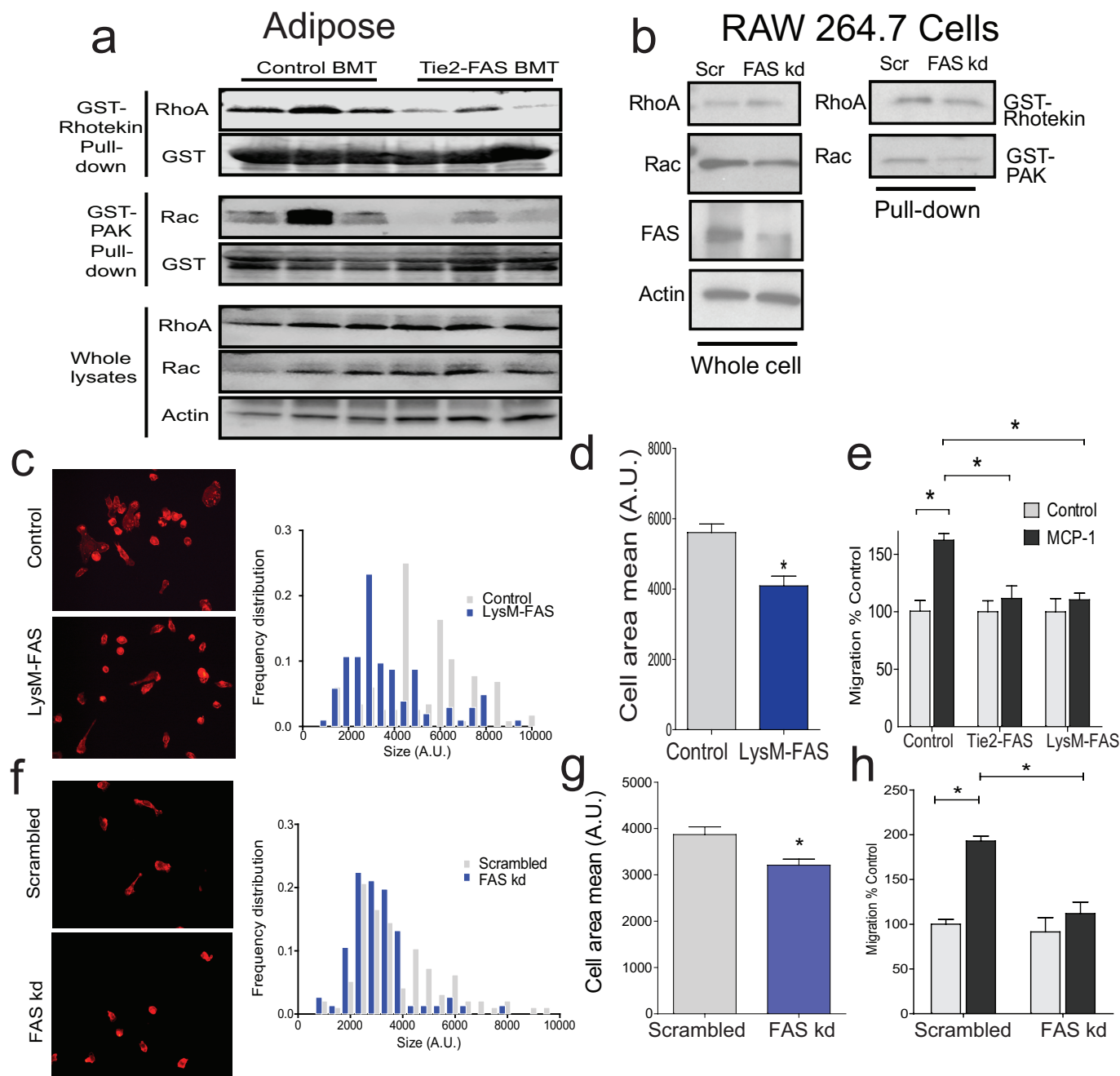
whole membranes ( $n = 8$ ). **d**, Quantification of the relative number of phosphatidylcholine, sphingomyelin, and phosphatidylethanolamine species by the detergent method ( $n = 8$  for control,  $n = 7$  for LysM–FAS). **e**, Relative number of lipid species by the non-detergent, high-pH carbonate buffer method ( $n = 4$ ). \* $P < 0.05$ , data are mean  $\pm$  s.e.m.





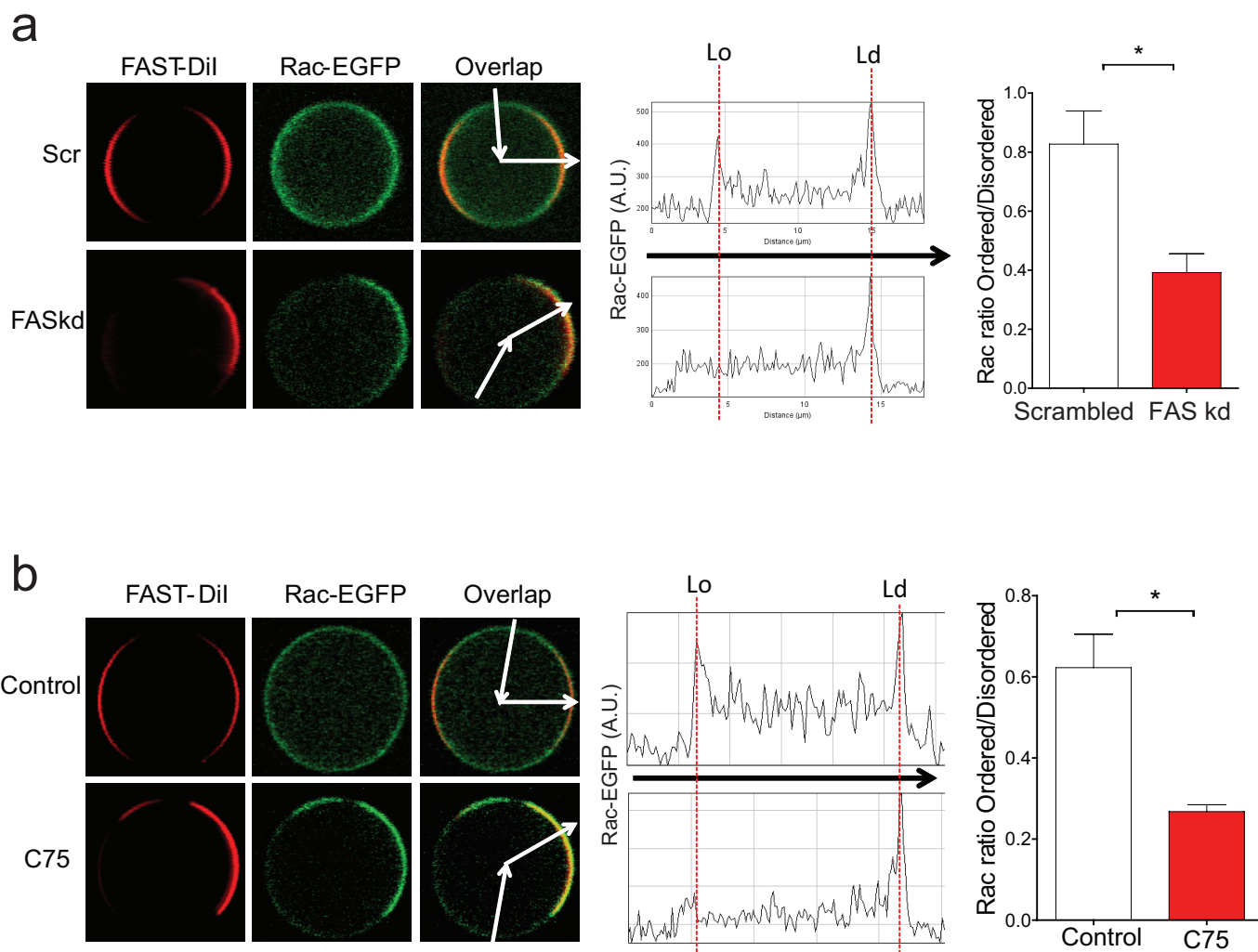
**Extended Data Figure 6 | FAS is required for compartmentalizing proteins in DRMs.** **a**, SILAC strategy for DRM proteomic analysis in bone-marrow-derived macrophages (BMDMs). **b**, Heat map of relative abundance for DRM proteins in both strains (LysM-FAS and Tie2-FAS) of FAS-deficient BMDMs as compared to FAS-replete BMDMs. **c**, SILAC strategy for whole-membrane proteomic analysis. **d**, Heat map of relative abundance for membrane-associated proteins in both FAS-deficient

models as compared to FAS-replete cells. **e**, SILAC quantification of DRM (top) and whole-membrane (bottom) proteomes in LysM-FAS and Tie2-FAS BMDMs relative to controls. **f**, **g**, DRM proteins affected by FAS deficiency (protein numbers in parentheses) in functional classes (**f**) and pathway analysis of affected proteins in macrophage inflammation (**g**). **h**, Heat map of select proteins from **e**. **i**, Western blot confirming heat-map changes of proteins indicated in red in **h**.



**Extended Data Figure 7 | Rho family GTPase activity and cell phenotyping in the setting of FAS deficiency.** **a**, Western blot for activated RhoA (top panel), activated Rac (middle panel) and total RhoA and Rac (bottom panel) in adipose tissue of control and Tie2-FAS BMT mice. **b**, Western blotting of whole cells (left), or cells with activated Rho GTPases (right with substrate pull-down) in scrambled (control) or FAS-knockdown RAW 264.7 cells. **c**, Cell spreading images of FAS-replete (control) and FAS-deficient (LysM-FAS) bone-marrow-derived macrophages (left) with corresponding frequency distribution of cell

spreading (right). **d**, Cell area for control ( $n = 116$ ) and LysM-FAS ( $n = 103$ ) cells. **e**, MCP1-induced Transwell migration of Tie2-FAS and LysM-FAS cells ( $n = 4$ ). **f**, Cell spreading images of scrambled (control) and FAS-deficient (FAS kd) RAW 264.7 cells (left) with the corresponding frequency distribution of cell spreading (right). **g**, Cell area for control ( $n = 108$ ) and FAS knockdown ( $n = 102$ ) RAW 264.7 cells. **h**, MCP1-induced Transwell migration of FAS-knockdown RAW 264.7 cells, expressed as a percentage of the value obtained for control cells ( $n = 4$ ).  $*P < 0.05$ , data are mean  $\pm$  s.e.m.

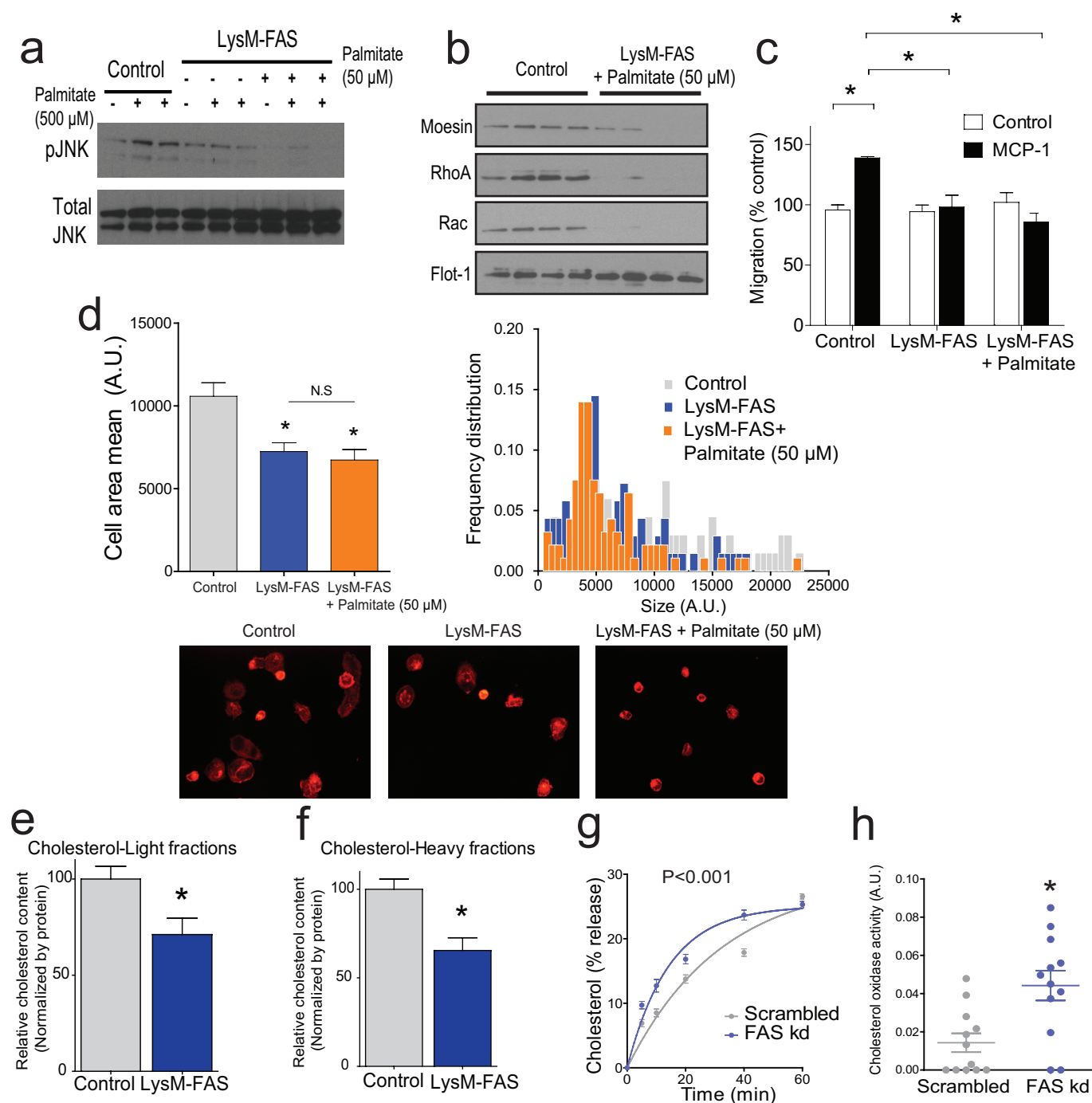


**Extended Data Figure 8 | Disruption of Rho GTPase membrane localization with genetic and pharmacologic inhibition of FAS.**

**a, b,** RAW 264.7 cells were transfected with a Rac-EGFP construct then subjected to FAS knockdown (**a**) or incubated with the nonspecific FAS inhibitor C75 (**b**), followed by preparation of GPMVs and phase imaging. Liquid-ordered and liquid-disordered domains for quantification are

indicated by arrows. Quantification of GFP fluorescence (with domains indicated) is shown in the middle and quantification of Rac partitioning in ordered versus disordered domains is shown on the right. Imaging analyses were repeated in 3 independent experiments and data shown are from representative experiments;  $n = 6$  vesicles per group.  $*P < 0.05$ , data are mean  $\pm$  s.e.m.

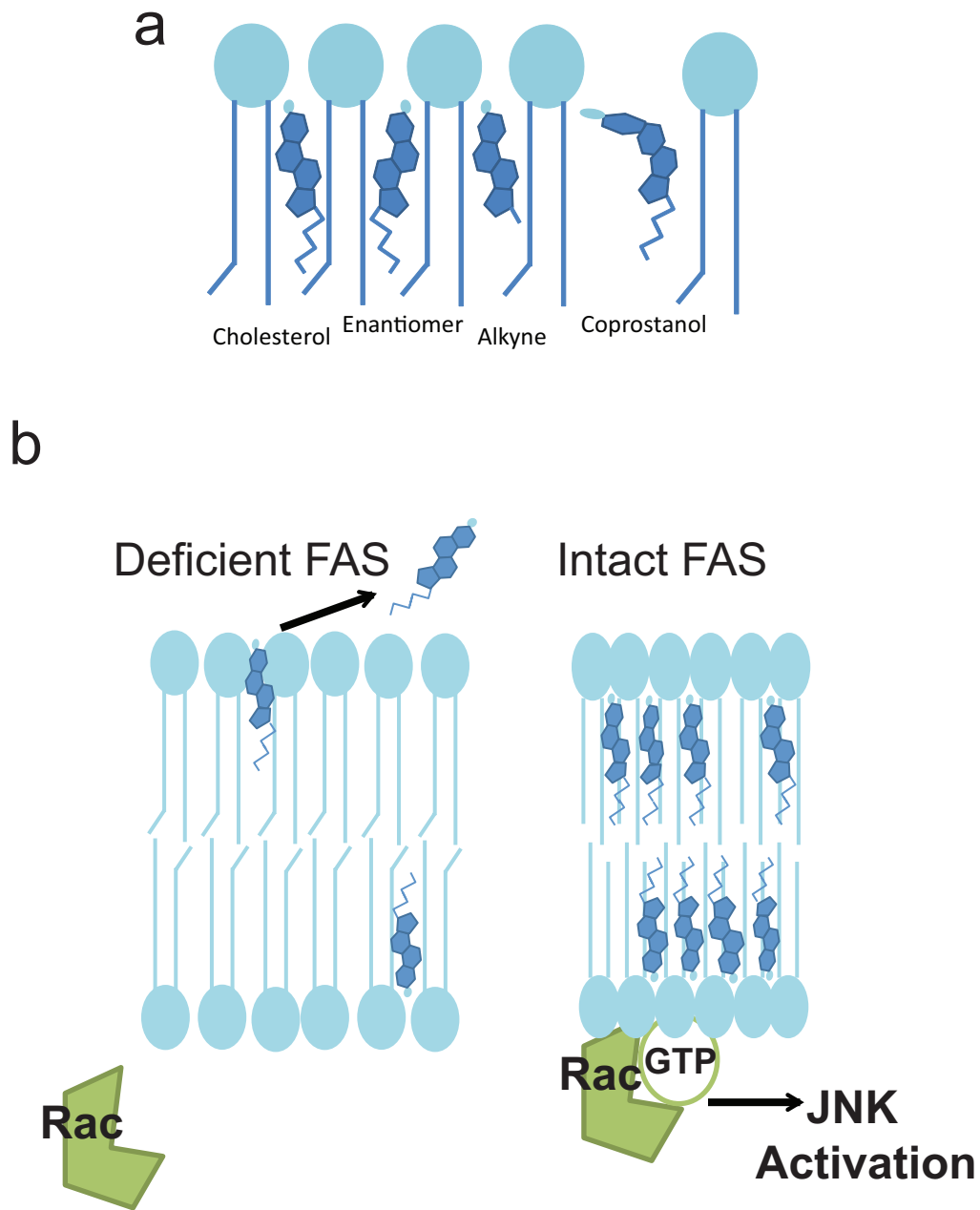




#### Extended Data Figure 9 | FAS-dependent macrophage response to chronic palmitate treatment and cholesterol characterization.

**a**, Changes in pJNK levels in response to acute exposure to high-dose palmitate (500  $\mu$ M) in control macrophages and in FAS-knockout bone-marrow-derived macrophages chronically pretreated with vehicle or 50  $\mu$ M palmitate for 24 h. **b**, Western blots of DRMs of LysM-FAS cells incubated with 50  $\mu$ M palmitate for 24 h. **c**, MCP1-induced Transwell migration assays of control and LysM-FAS cells, demonstrating that palmitate incubation does not rescue the defect in macrophage motility with FAS deficiency ( $n = 4$ ). **d**, Cell spreading assays for control macrophages, FAS-deficient macrophages, and FAS-deficient macrophages treated with

50  $\mu$ M palmitate for 24 h. Cell area quantification is shown on the left, with the frequency distribution of cell spreading on the right, and images of macrophage spreading on the bottom. For the frequency distribution of cell spreading,  $n = 108$  for control,  $n = 101$  for LysM-FAS and  $n = 105$  for LysM-FAS with palmitate. **e**, **f**, Cholesterol content of fractions from LysM-FAS cells ( $n = 4$ ). **g**, Cholesterol release to methyl- $\beta$ -cyclodextrin in control and FAS-knockdown RAW 264.7 cells ( $n = 4$ ). **h**, Cholesterol oxidase activity in control and FAS knockdown RAW 264.7 cells ( $n = 12$ ). \* $P < 0.05$ , except in **g** where  $P < 0.001$  by nonlinear curve-fit comparison; data are mean  $\pm$  s.e.m.



**Extended Data Figure 10 | Schematic rendering of sterols in membranes.** **a**, Generalized planar and non-planar structures of the sterols used to load macrophages. **b**, Diagram of FAS-dependent cholesterol retention at the plasma membrane leading to JNK activation. In the FAS-deficient state, the phospholipid environment is characterized

by greater levels of unsaturated fatty acids whereas in the FAS-replete state, phospholipids are enriched for saturated fatty acids. Phospholipids are depicted schematically and not intended to represent authentic unsaturated fatty acid structures.

# Reconstitution *in vitro* of the entire cycle of the mouse female germ line

Orie Hikabe<sup>1\*</sup>, Nobuhiko Hamazaki<sup>1</sup>, Go Nagamatsu<sup>1</sup>, Yayoi Obata<sup>2</sup>, Yuji Hirao<sup>3</sup>, Norio Hamada<sup>1,4</sup>, So Shimamoto<sup>1</sup>, Takuya Imamura<sup>1</sup>, Kinichi Nakashima<sup>1</sup>, Mitinori Saitou<sup>5,6,7,8</sup> & Katsuhiko Hayashi<sup>1,9\*</sup>

**The female germ line undergoes a unique sequence of differentiation processes that confers totipotency to the egg<sup>1,2</sup>. The reconstitution of these events *in vitro* using pluripotent stem cells is a key achievement in reproductive biology and regenerative medicine. Here we report successful reconstitution *in vitro* of the entire process of oogenesis from mouse pluripotent stem cells. Fully potent mature oocytes were generated in culture from embryonic stem cells and from induced pluripotent stem cells derived from both embryonic fibroblasts and adult tail tip fibroblasts. Moreover, pluripotent stem cell lines were re-derived from the eggs that were generated *in vitro*, thereby reconstituting the full female germline cycle in a dish. This culture system will provide a platform for elucidating the molecular mechanisms underlying totipotency and the production of oocytes of other mammalian species in culture.**

One of the key goals in developmental and reproductive biology is to reconstitute the entire process of gametogenesis in culture. Specifically, owing to its biological significance, reconstitution of oogenesis using pluripotent stem cells that yields functional eggs has long been sought<sup>3–5</sup>. Eggs originate from primordial germ cells (PGCs), which are specified at around embryonic day 6.5 (E6.5) in mice<sup>6</sup>. PGCs then migrate into the gonads, enter meiosis in female embryos<sup>7</sup>, and therefore become primary oocytes. Following puberty, primary oocytes begin to grow to mature oocytes that are fully ready for fertilization. We previously reported a culture system in which mouse pluripotent stem cells differentiated into PGC-like cells (PGCLCs) that were capable of differentiating into functional oocytes by transplantation into immunocompromised female mice<sup>8</sup>. However, this system recapitulates a quite short window of embryogenesis, from the blastocyst to migratory PGC stage, requiring only 4 to 5 days *in vivo*. There thus remains a long period of time (perhaps over a month) that must be reconstituted to produce mature oocytes in culture.

To reconstitute the entire process of oogenesis *in vitro*, the culture period was divided into three sections in this study: (1) *in vitro* differentiation (IVDi), (2) *in vitro* growth (IVG) and (3) *in vitro* maturation (IVM), in which oogenesis would proceed to primary oocytes in the secondary follicle, fully grown germinal vesicle oocytes and metaphase II (MII) oocytes, respectively (Fig. 1a). As we reported previously<sup>8</sup>, PGCLCs differentiated from a female embryonic stem cell (ESC) line, BVSC18 (129<sup>+/Ter</sup>/svj (agouti) × C57BL/6), that harbours *Blimp1-mVenus* (BV) and *stella-ECFP* (SC) reporter constructs, were aggregated with E12.5 female gonadal somatic cells of albino ICR strain. In IVDi, the aggregates, hereafter called reconstituted ovaries (rOvaries), were placed on Transwell-COL, a culture insert coated with collagen, and soaked with  $\alpha$ MEM-based medium. At 4 days of culture, the medium was replaced with StemPro34-based medium, as

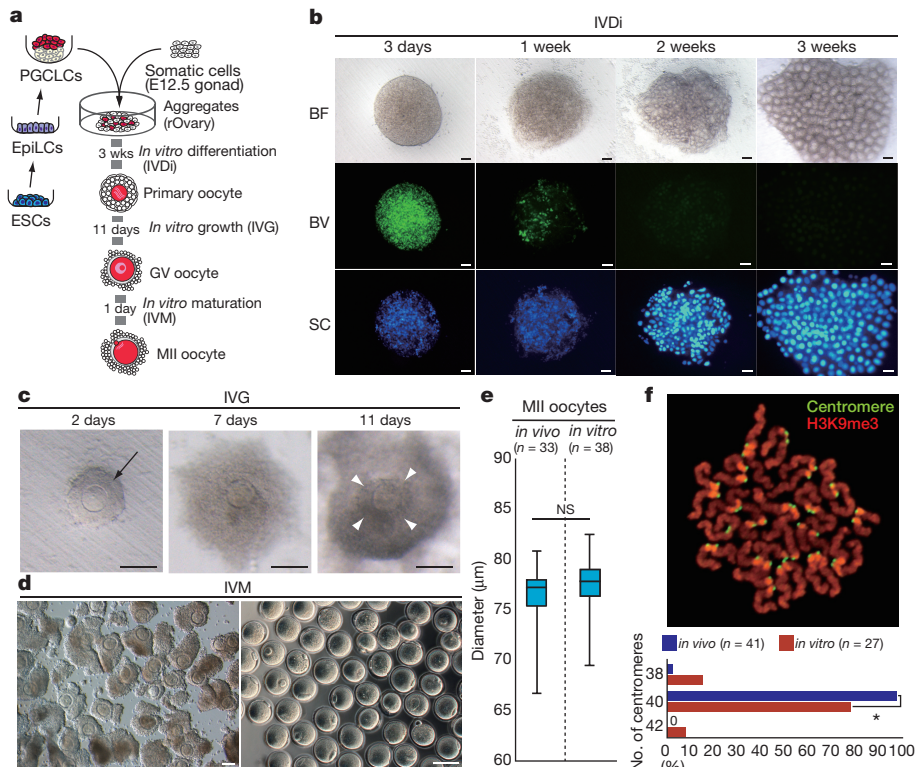
the combination of media yielded a high number of primary oocytes with a follicle structure in preliminary culture experiments (Extended Data Fig. 1a). To prevent multiple oocyte follicle formation such as that seen frequently in culture (Extended Data Fig. 1b), we added the oestrogen inhibitor ICI182780 to the culture<sup>9</sup>. The rOvaries were filled throughout with follicle structures, each of which possesses a single oocyte, in an ICI182780-dependent manner (Extended Data Fig. 1b, c). During IVDi culture, BV, a marker of early PGCs<sup>10</sup>, was detectable at 3 days of culture, but became weak after one week of culture (Fig. 1b). At two weeks of culture, BV disappeared and SC—a marker of both oocytes and PGCs—became prominent in rOvaries, and at three weeks of culture, a number of SC-positive primary oocytes were observed in rOvaries (Fig. 1b). The close observation of IVDi revealed that clusters of PGCLCs were formed by 5 days of culture and gradually fragmented from 5 to 9 days, after which, follicles were formed around 11 days of culture (Extended Data Fig. 2a). FoxL2, a functional marker of granulosa cells, was detectable in cells surrounding the oocytes at 21 days of culture (Extended Data Fig. 2a). Immunofluorescence analysis of the meiotic chromosome revealed the robust progression of meiotic prophase I from 5 to 9 days of culture (Extended Data Fig. 2b, c). PGCLCs before aggregation are equivalent to E9.5 (refs 8, 11). Therefore, the meiotic prophase I in PGCLCs progresses during the period corresponding to E14.5 (E9.5 + 5 days) through to E18.5 (E9.5 + 9 days), which is also seen in meiotic prophase I in fetal ovaries *in vivo*<sup>12</sup>. At the pachytene stage, however, asynapsis between homologous chromosomes occurred more frequently *in vitro* (53.8%) than *in vivo* (5.3%) (Extended Data Fig. 2d). This was partially due to culture *in vitro*, as the percentage increased even in primary oocytes in the organ culture of gonads under the same condition (Extended Data Fig. 2d). Although chromosome pairing *in vitro* was not as accurate as that *in vivo*, IVDi culture yielded a large number of secondary follicle-like structures (2FLs) at three weeks of culture; on average ( $n = 8$ ),  $237.3 \pm 27.3$  2FLs were formed per rOvary. SC-negative residual oocytes were sparsely scattered in the rOvaries (Fig. 1b, Extended Data Fig. 1c); the percentage of SC-negative oocytes was  $4.4 \pm 2.1\%$  (the mean  $\pm$  s.d.,  $n = 4$ , 3163 oocytes in total). These results demonstrated that a robust number of primary oocytes were induced from ESCs under the culture conditions employed.

Oocyte growth accompanied with follicular growth is tightly controlled by gonadotropins<sup>2</sup>. Follicle-stimulating hormone plays a central role in the proliferation and maturation of granulosa cells. In culture with IVG medium that contains follicle-stimulating hormone, however, proliferation of the granulosa cell layer was limited in 2FLs located at the edge of the rOvary (Extended Data Fig. 3a), suggesting that 2FLs in the centre portion lack signalling and/or space for cell growth. Therefore, individual 2FLs were manually separated from the

<sup>1</sup>Department of Stem Cell Biology and Medicine, Graduate School of Medical Sciences, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582, Japan. <sup>2</sup>Department of Bioscience, Tokyo University of Agriculture, 1-1-1, Sakuragaoka, Setagaya-ku, Tokyo 156-8502, Japan. <sup>3</sup>NARO Institute of Livestock and Grassland Science, Ikenodai 2, Tsukuba 305-0901, Japan. <sup>4</sup>Department of Obstetrics and Gynecology, Graduate School of Medical Sciences, Kyushu University, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582, Japan. <sup>5</sup>Department of Anatomy and Cell Biology, Graduate School of Medicine, Kyoto University, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan. <sup>6</sup>Center for iPS Cell Research and Application, Kyoto University, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan. <sup>7</sup>Institute for Integrated Cell-Material Sciences, Kyoto University, Yoshida-Ushinomiya-cho, Sakyo-ku, Kyoto 606-8501, Japan. <sup>8</sup>JST, ERATO, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan. <sup>9</sup>JST, PRESTO, Maidashi 3-1-1, Higashi-ku, Fukuoka 812-8582, Japan.

\*These authors contributed equally to this work.





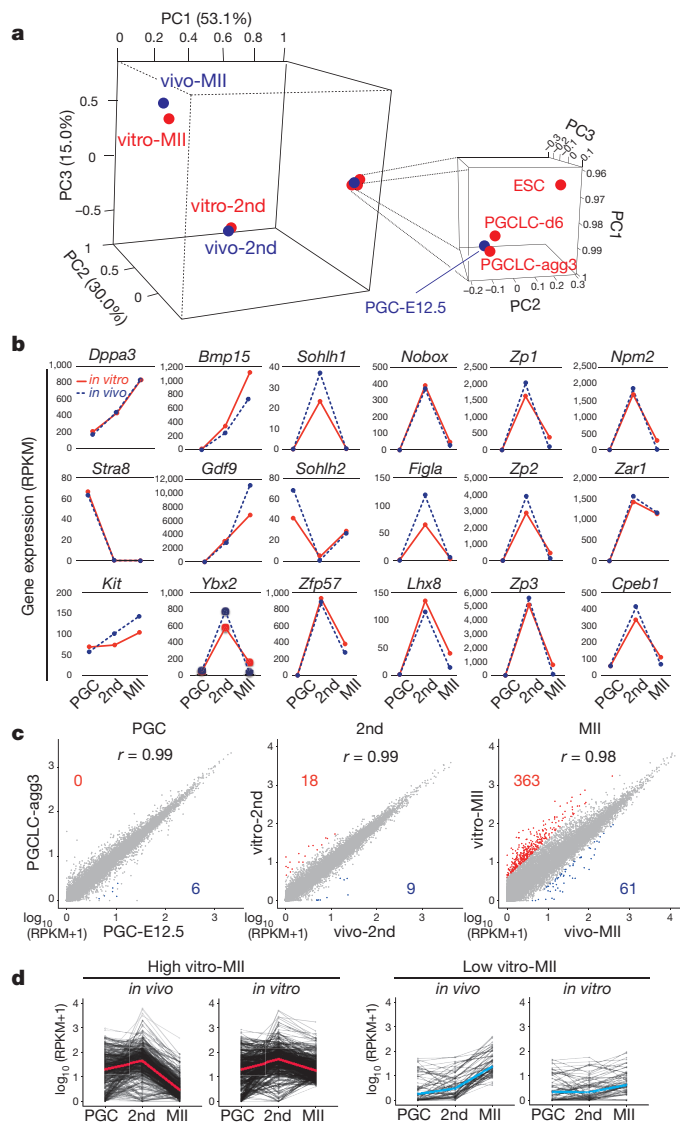
**Figure 1 | Oocyte production from pluripotent stem cells.** **a**, A schematic of oocyte production *in vitro*. **b**, rOvary during IVDi. A representative rOvary on days indicated are shown. BF, bright field; BV, *Blimp1-mVenus*; SC, *stella-ECFP*. **c**, Follicular growth during IVG. Granulosa cells (black arrow) proliferated and then formed a cumulus-oocyte complex (white arrowheads). **d**, Cumulus-oocyte complexes before IVM (left) and MII oocytes after IVM (right). A large image of the MII oocytes is shown in Extended Data Fig. 3d. **e**, Diameters of the *in-vitro*-generated MII oocytes

rOvaries (Extended Data Fig. 3b). As expected, granulosa cells proliferated and cumulus-oocyte complexes were formed in the isolated 2FLs (Fig. 1c). When cultured for 11 days, primary oocytes in 2FLs grew to germinal vesicle oocytes (Fig. 1c). In total, 3,198 fully grown germinal vesicle oocytes were obtained from 58 rOvaries in three culture experiments (Extended Data Fig. 3c); on average, 55.1 fully grown oocytes were obtained from one rOvary. When transferring them under the IVM culture conditions, 28.9% of the germinal vesicle oocytes extruded a 1st polar body (Fig. 1d and Extended Data Fig. 3c, d). The percentage of SC-negative MII oocytes was  $4.6 \pm 0.7\%$  (mean  $\pm$  s.d.,  $n = 3,565$  MII oocytes in total), suggesting that the contamination rate of endogenous oocytes remained low. The diameters of SC-positive MII oocytes generated *in vitro* were comparable to those of MII oocytes *in vivo* (Fig. 1e). Consistently, the maternal pattern of DNA methylation in the differentially methylated region (DMR) of imprinted genes, which is established in an oocyte size-specific manner<sup>13</sup>, was almost complete at the representative gene loci tested (*H19* and *Igf2r*) (Extended Data Fig. 3e). The frequency of aneuploidy increased to some extent, but most (77.8%) of the MII oocytes generated *in vitro* had a correct number of chromosomes (Fig. 1f). These results demonstrate that, despite sporadic flaws in some oocytes, a robust number of MII oocytes were produced from pluripotent stem cells in a dish.

To evaluate oogenesis *in vitro*, the global transcription dynamics during the culture were determined by RNA sequencing (RNA-seq) analysis. We used poly(A) RNA from ESCs, PGCLCs before aggregation (PGCLCs-d6), PGCLCs at 3 days of aggregation (PGCLC-agg3), SC-positive primary oocytes in 2FLs (vitro-2nd) and MII oocytes generated *in vitro* (vitro-MII). For comparison, poly(A) RNAs from E12.5 female PGCs (PGCs-E12.5), primary oocytes in secondary follicles in post-natal day 8 (P8) ovaries (vivo-2nd) and MII oocytes

in 2 independent experiments. NS, not significant ( $t$ -test,  $P > 0.05$ ). **f**, Meiotic chromosome in *in-vitro*-generated MII oocytes in 3 independent experiments. Centromeres (green) and H3K9me3 (red) (upper), and the number of centromeres in the MII oocytes (bottom) are shown. The percentage of *in-vitro*-generated MII oocytes harbouring 40 centromeres was lower (chi-squared test,  $*P < 0.05$ ) than that of MII oocytes *in vivo*. Scale bars in **b**, **c** and **d**, 100  $\mu$ m.

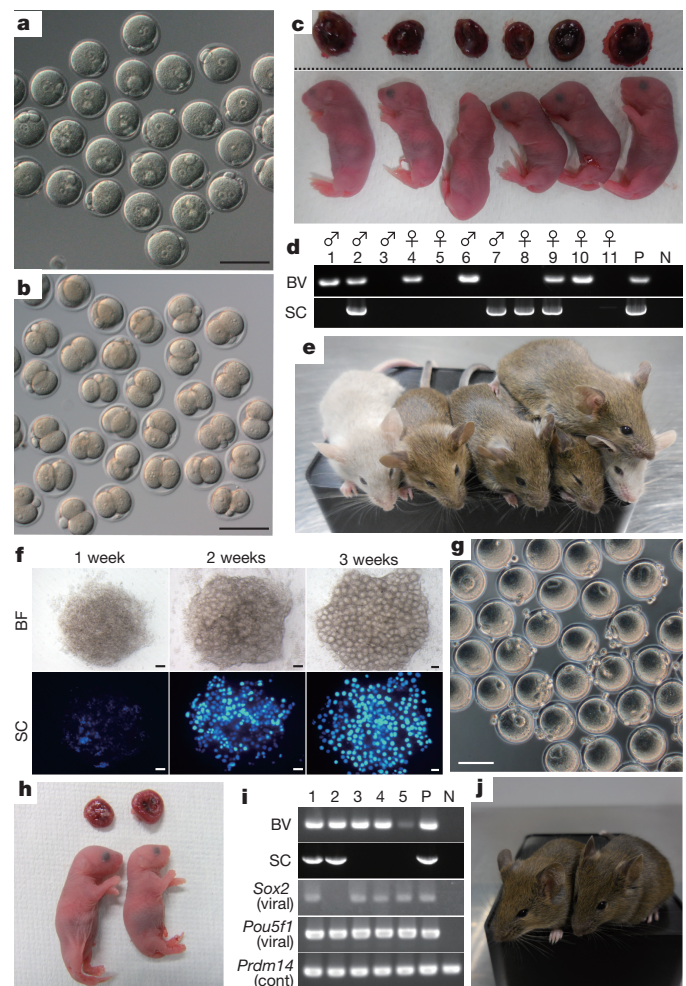
superovulated from adult ovaries (vivo-MII) were used. Principal component analysis showed that oogenesis *in vitro* largely recapitulated the differentiation process *in vivo* (Fig. 2a). Consistently, the expression dynamics of genes involved in oogenesis was similar to those *in vivo* (Fig. 2b). Repetitive elements, such as long interspersed elements, short interspersed elements and satellite sequences, were repressed in oocytes both *in vitro* and *in vivo*, and the transcripts of long terminal repeat (LTR) transposons was temporarily upregulated in the primary oocytes both *in vivo* and *in vitro* (Extended Data Fig. 4a). Endogenous retrovirus group K (ERVK) and endogenous retrovirus-like and mammalian apparent LTR-retrotransposon (ERVL-MaLR) accounted for most of the abundant LTR transposon transcripts (Extended Data Fig. 4b). Moreover, the retrotransposons LTR 10 (RLTR10) and mouse transcript A (MTA) transcripts were enriched in ERVK and ERVL-MaLR, respectively (Extended Data Fig. 4c, d). These observations are consistent with previous reports that the transcripts of transposons enriched in growing oocytes are tightly related to oocyte-specific transcriptional regulation<sup>14,15</sup>. Our results therefore indicate that such an oocyte-specific regulation was firmly reconstituted during oogenesis in culture. The correlation coefficient between global transcripts *in vivo* and *in vitro* at each oogenesis stage was more than 0.98 (Fig. 2c). However, careful comparison of the transcriptome identified differentially expressed genes (DEGs) between oogenesis *in vivo* and *in vitro*. The numbers of DEGs with >4-fold change were 6 and 27 at the PGC- and primary-oocyte-stage, respectively, which was a relatively subtle difference (Fig. 2c). In contrast, at the MII-oocyte-stage, 424 genes were detected as the DEGs; 363 and 61 genes were up- and downregulated, respectively (Fig. 2c). Of note, the DEGs upregulated in MII oocytes *in vitro* were the genes whose transcripts were reduced from primary oocytes to MII oocytes *in vivo*,



**Figure 2 | Transcriptome analysis of oogenesis *in vitro*.** **a**, Principal component analysis of each developmental stage. **b**, Gene expression dynamics during oogenesis *in vitro*. RPKMs of gene products in oogenesis *in vitro* (red) and *in vivo* (blue) are shown. PGC on the x axis indicate the comparisons between PGCLC-aggr3 and PGCs-E12.5. **c**, Scatter plot comparison of transcripts between each stage of oogenesis *in vitro* and *in vivo*. Pearson correlation coefficients ( $r$ ) are shown in the plot. Genes exhibiting more than fourfold higher (red) or fourfold lower (blue) expression *in vitro* than *in vivo* are shown. **d**, Expression dynamics of the DEGs in the MII oocyte stage. The red and blue lines indicate the median values. All expression data are based on the mean values from 3 biological replicates, except for PGCLC-d6 that is from 2 biological replicates.

and vice versa: the DEGs downregulated in MII oocytes *in vitro* were the genes whose transcripts were increased from primary oocytes to MII oocytes *in vivo* (Fig. 2d). This indicates that oocyte growth during IVG and IVM culture was compromised in a subset, or perhaps all, of the oocytes. Interestingly, gene ontology analysis illustrated that DEGs upregulated in MII oocytes *in vitro* were related to mitochondrial function (Extended Data Fig. 4e and Supplementary Table 1). Misregulation of these genes might attenuate the potential of MII oocytes generated *in vitro*. Analysis of the metabolic pathway will thus provide information about the refinement of the IVG and IVM culture.

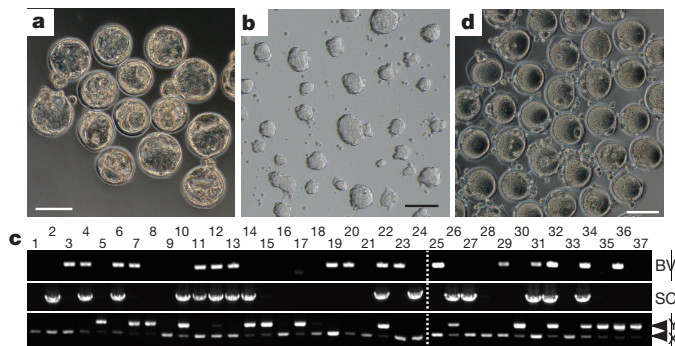
Whether the MII oocytes are capable of developing to offspring is the most stringent criterion for evaluation of the culture system. To test this capability, *in-vitro*-generated MII oocytes were subjected to *in vitro* fertilization (IVF) with wild-type sperm of albino ICR strain.



**Figure 3 | Offspring from *in-vitro*-generated eggs.** **a**, Fertilized eggs (**a**) and 2-cell embryos (**b**) from IVF using BVSC18 ESC-derived MII oocytes. **c**, 6 representative pups and their placentas from BVSC18 ESC-derived MII oocytes. **d**, Genotyping by the BV and SC transgenes. **e**, The 6 mice at 4 weeks after birth. **f**, IVDi using TTF4FC6. **g**, MII oocytes from TTF4FC6 after IVM culture. **h**, Two representative pups and their placentas from TTF4FC6-derived MII oocytes. **i**, Genotyping of the pups from TTF4FC6. PCR products from the BV and SC transgenes, exogenous retroviral *Sox2* and *Pou5f1* and the *Prdm14* locus as a positive control for the PCR reaction are shown. **j**, Two mice (from **h**) at 4 weeks after birth. Scale bars, 100  $\mu$ m. P, positive control (genomic DNA of a BVSC mouse in **d** and of TTF4FC6 in **i**). N, negative control (genomic DNA of a wild-type mouse). For gel source data see Supplementary Fig. 1.

By IVF, *in-vitro*-generated oocytes were fertilized and developed to 2-cell embryos (Fig. 3a, b). 11 (3.5%) out of 316 2-cell embryos transferred to pseudopregnant ICR females were successfully delivered as viable pups with coloured eyes (Fig. 3c, Extended Data Fig. 5a, 5b), showing that the pups were from BVSC18 ESC-derived oocytes, but not from ICR oocytes that could mingle in gonadal somatic cells. The reporter constructs were largely segregated in the expected Mendelian manner (Fig. 3d). The placentas of the pups were heavier than those of wild-type mice (Extended Data Fig. 5c), but this phenomenon was less pronounced than the placentomegaly observed in clone mice<sup>16</sup>. In addition, the pups tended to be slightly heavier than the wild-type mice (Extended Data Fig. 5c), which may have been due to the small number (either one or two) of pups in the uterine horn on Caesarean section. Using 2-cell embryos from *in-vitro*-generated oocytes, the success rate of full-term development was 3.5%, which is much lower than that from *in-vivo*-generated oocytes (61.7%, Extended Data Fig. 6a). A comparison of developmental progression from genetically matched oocytes *in vivo* (129X1/svj  $\times$  C57BL/6) demonstrates that the





**Figure 4 | Reconstitution of the second cycle of oogenesis *in vitro*.** **a, b,** Derivation of rESCs. Blastocysts from *in-vitro*-generated eggs (**a**) were used for derivation of rESC (**b**). **c,** Genotyping of rESCs. In total, 37 rESC lines were derived in 2 independent experiments. The BVSC constructs and sex chromosome (X or Y) were determined by PCR. **d,** MII oocytes from rESCs. Scale bars, 100  $\mu$ m. For gel source data see Supplementary Fig. 1.

fertilization rates were not significantly different, but embryogenesis was frequently retarded at various stages, such as the cleavage stage, and early and late gestation (Extended Data Fig. 6b–e). The developmental arrest may have been at least partly attributable to aneuploidy (Fig. 1f) and aberrant gene expression (Fig. 2c, d). Nevertheless, the successful production of pups demonstrates that the culture system fulfills the condition of producing functional eggs. Importantly, all of the obtained pups grew up normally without evidence of premature death (Fig. 3e, Extended Data Fig. 5c). Combined bisulfite restriction analysis (COBRA) and bisulfite sequencing of the DMRs revealed that the epigenetic status of the imprinted loci was maintained in a manner comparable to that of wild-type mice (Extended Data Fig. 5d, e). Both females and males from *in-vitro*-generated oocytes were fertile (Extended Data Fig. 5f) and remained alive for at least 11 months without any apparent abnormality (Extended Data Fig. 5g).

To verify the robustness and reproducibility of the culture system, we induced oogenesis in five cell lines *in vitro*: BVSC14, which is another ESC line of the same mouse strain as BVSC18, two induced pluripotent stem cell (iPSC) lines from mouse embryonic fibroblasts (MEFs) and two iPSC lines from 10-week-old adult tail tip fibroblasts (TTFs). BVSC14 and the two MEF-derived iPSC lines (MEF4FRC9 and MEF4FC14) gave rise to a number of oocytes (Extended Data Fig. 7a, Extended Data Table 1). They contained fully potent oocytes, as pups with coloured eyes were obtained from oocytes from each cell line (Extended Data Figs 7a, 8a, b, Extended Data Table 1). The weights of their placentas and bodies showed a similar trend as seen in BVSC18 (Extended Data Fig. 7b). The reporter genes and retrovirus-derived exogenous genes for the reprogramming were detectable in the pups, as expected (Extended Data Figs 7c, d, 8a, b). All of the pups grew up normally without evidence of premature death. Notably, fully potent oocytes were also generated from the two TTF-derived iPSC lines (TTF4FC6 and TTF4FRC3) (Fig. 3f, g, Extended Data Table 1). Upon fertilization with sperm from ICR males followed by transfer to surrogate mothers, they gave rise to healthy pups with the retroviral genes and the reporter constructs or coloured eyes (Fig. 3h, i, Extended Data Fig. 7d, 7e, 8c, Extended Data Table 1). The weights of the placentas were varied, and some of them were comparable to wild-type placenta (Extended Data Fig. 7b). Although 2 pups were cannibalized by the nursing mother at the day of birth, the remaining 6 pups grew without any apparent abnormality such as tumour formation or developmental arrest (Fig. 3j). Both females and males from iPSC-derived oocytes were fertile (Extended Data Fig. 8d).

Finally, blastocysts from *in-vitro*-generated oocytes were used to derive ESCs, thereby realizing the reconstitution of an entire cycle of the female germ line. The 51 blastocysts from BVSC18-derived oocytes were placed on embryonic fibroblasts in culture medium containing

CHIR99021, PD0325901 and leukaemia inhibitory factor (2i+LIF)<sup>17</sup>. 72.5% of the blastocysts (37 out of 51) gave rise to ESC lines (Fig. 4a–c), which were designated regenerated ESCs (rESCs). The BVSC reporter genes and sex chromosomes were segregated to each rESC clone in the expected Mendelian manner (Fig. 4c). When the rESCs (clone 6) were injected into wild-type blastocysts, they contributed to multiple tissues, including PGCs of the chimaera embryos (Extended Data Fig. 9a, b). As the success rate of rESC derivation was higher than that of full-term development, it is possible that some rESC lines are not fully competent. A critical comparison of pluripotency, gene expression and epigenetic status between these rESC lines might provide important clues to clarify the heterogeneity in quality of *in-vitro*-generated oocytes. We successfully produced rESC-derived MII oocytes that were undergoing second rounds of meiosis from the initial ESCs (Fig. 4d, Extended Data Fig. 9c). Taken together, these results demonstrated the successful reconstitution of an entire cycle of the female germ line *in vitro* (Extended Data Fig. 9d). One limitation of the reconstitution, however, is that the culture system used gonadal somatic cells that were obtained from the embryos. Requirement of gonadal somatic cells would be a critical point, if human PGCLCs<sup>18,19</sup> are to be progressed to later stages of gametogenesis. One possible way to overcome this issue would be to derive the gonadal somatic cell-like cells from pluripotent stem cells. Nevertheless, the culture system used here will provide an important platform for analysing the gene functions underlying oogenesis in addition to providing clues to the development of a similar culture system in other species.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 22 October 2015; accepted 22 September 2016.**

**Published online 17 October 2016.**

- Surani, M. A., Hayashi, K. & Hajkova, P. Genetic and epigenetic regulators of pluripotency. *Cell* **128**, 747–762 (2007).
- Edson, M. A., Nagaraja, A. K. & Matzuk, M. M. The mammalian ovary from genesis to revelation. *Endocr. Rev.* **30**, 624–712 (2009).
- Hübner, K. *et al.* Derivation of oocytes from mouse embryonic stem cells. *Science* **300**, 1251–1256 (2003).
- Qing, T. *et al.* Induction of oocyte-like cells from mouse embryonic stem cells by co-culture with ovarian granulosa cells. *Differentiation* **75**, 902–911 (2007).
- Salvador, L. M., Silva, C. P., Kostetskii, I., Radice, G. L. & Strauss, J. F., III. The promoter of the oocyte-specific gene, *Gdf9*, is active in population of cultured mouse embryonic stem cells with an oocyte-like phenotype. *Methods* **45**, 172–181 (2008).
- Ohinata, Y. *et al.* Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**, 207–213 (2005).
- McLaren, A. & Southee, D. Entry of mouse embryonic germ cells into meiosis. *Dev. Biol.* **187**, 107–113 (1997).
- Hayashi, K. *et al.* Offspring from oocytes derived from *in vitro* primordial germ cell-like cells in mice. *Science* **338**, 971–975 (2012).
- Morohaku, K. *et al.* Complete *in vitro* generation of fertile oocytes from mouse primordial germ cells. *Proc. Natl Acad. Sci. USA* **113**, 9021–9026 (2016).
- Ohinata, Y., Sano, M., Shigeta, M., Yamanaka, K. & Saitou, M. A comprehensive, non-invasive visualization of primordial germ cell development in mice by the *Prdm1*-mVenus and *Dppa3*-ECFP double transgenic reporter. *Reproduction* **136**, 503–514 (2008).
- Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519–532 (2011).
- McClellan, K. A., Gosden, R. & Taketo, T. Continuous loss of oocytes throughout meiotic prophase in the normal mouse ovary. *Dev. Biol.* **258**, 334–348 (2003).
- Lucifero, D., Mann, M. R., Bartolomei, M. S. & Trasler, J. M. Gene-specific timing and epigenetic memory in oocyte imprinting. *Hum. Mol. Genet.* **13**, 839–849 (2004).
- Peaston, A. E. *et al.* Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **7**, 597–606 (2004).
- Veselovska, L. *et al.* Deep sequencing and de novo assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. *Genome Biol.* **16**, 209 (2015).
- Tanaka, S. *et al.* Placentomegaly in cloned mouse concepti caused by expansion of the spongiotrophoblast layer. *Biol. Reprod.* **65**, 1813–1821 (2001).



17. Ying, Q. L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
18. Irie, N. *et al.* SOX17 is a critical specifier of human primordial germ cell fate. *Cell* **160**, 253–268 (2015).
19. Sasaki, K. *et al.* Robust *in vitro* induction of human germ cell fate from pluripotent stem cells. *Cell Stem Cell* **17**, 178–194 (2015).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank Y. Takada for technical support, H. Ohta for technical advice on the iPSCs and embryo transfer, Y. Ohkawa for technical assistance on the RNA-seq, K. Kitajima and C. Meno for providing microscopes, F. Arai for providing FACS AriaII, and H. Leitch for proofreading the manuscript. We also thank the Research Support Center, Research Center for Human Disease Modeling, Kyushu University Graduate School of Medical Sciences for technical assistance. N.H. was a JSPS Research Fellow. This study was supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (KAKENHI #25114006); by JST-PRESTO; by the Uehara Memorial Foundation; and by the Takeda Science Foundation.

**Author Contributions** O.H., N.H., S.S. and K.H. performed the culture, embryo transfer, immunofluorescence, PCR analysis and chimera analysis; O.H., N.H., S.S. and K.H. performed the differentiation to mature oocytes and are able to replicate the entire process, independently. N.H., T.I., and K.N. performed the RNA-seq analysis. G.N. generated the iPSC lines. Y.O. and Y.H. conducted the assessment of the culture conditions. M.S. contributed to the inception of the study. K.H. designed the experiments and wrote the manuscript.

**Author Information** The RNA-seq data have been deposited at Gene Expression Omnibus (GEO) database under accession number GSE79729. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.H. ([hayashik@hgs.med.kyushu-u.ac.jp](mailto:hayashik@hgs.med.kyushu-u.ac.jp)).

**Reviewer Information** *Nature* thanks D. Egli, S. Mitalipov and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

A detailed protocol for oocyte production from pluripotent stem cells is available on the Nature Protocol Exchange at <http://dx.doi.org/10.1038/protex.2016.065>.

**Data reporting.** No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

**Animals.** All animal experiments were performed under the ethical guidelines of Kyushu University (the approval numbers A26-260-0 and A28-109-0). ICR, C57BL/6J and 129X1/svj mice were purchased from Japan SLC. 129X1/svj females were mated with C57BL/6J males to obtain hybrid mice (129X1/svj  $\times$  C57BL/6J). Pregnant females were killed by cervical dislocation to obtain E12.5 embryos. From female embryos, the gonads were isolated and then dissociated for rOvary formation<sup>20</sup>.

**ESCs and iPSCs.** Two independent ESC lines bearing the BVSC reporter constructs, BVSC18 and BVSC14, were used in this study. Both ESC lines were established in our laboratory from the blastocysts collected from independent pairs of 129<sup>+Ter</sup>/svj (agouti) females and C57BL/6 BVSC males<sup>8</sup>. These ESC lines were maintained under a 2i+LIF condition without feeders<sup>17</sup>.

For generation of iPSCs, MEFs and TTFs were obtained from E12.5 female embryos and 10-week-old female mice, respectively. Both embryos and mice were derived from mating between 129X1/svj (chinchilla) female and C57BL/6 BVSC males. Preparation of MEFs and TTFs and establishment of iPSCs in our laboratory were performed as described previously<sup>21,22</sup> with slight modifications. Briefly, for MEF preparation, the embryos were minced, digested by trypsin-EDTA (Invitrogen), and then cultured in DMEM containing 10% FBS supplemented with 2 mM glutamine and 1  $\times$  penicillin/streptomycin (MEF medium). For TTF preparation, the tails were cut, their epidermises were removed and the remaining tails were minced into about 1 cm pieces. They were then cultured with MF-start medium (Toyobo) for 5 days. The cells migrating out from the tail were cultured in MEF medium.

MEFs and TTFs were reprogrammed to iPSCs by introducing retroviral vectors containing the four genes (pMXs-*Pou5f1*, pMXs-*Sox2*, pMXs-*Klf4* and pMXs-*c-Myc*). MEFs and TTFs were cultured in MEF medium for four days and then were reseeded on STO feeder cells. One day after the reseed, the medium was changed to Knockout DMEM (Invitrogen) supplemented with 15% KnockOut Serum Replacement (KSR; Invitrogen), 2 mM glutamine, 1 mM non-essential amino acids, 55  $\mu$ M 2-mercaptoethanol, 10<sup>3</sup> U ml<sup>-1</sup> leukaemia inhibitory factor (LIF), and 1  $\times$  penicillin/streptomycin. The medium was changed every other day. When ES-like colonies emerged, which usually took more than 14 days after infection, the medium was changed to 2i+LIF. ES-like colonies were picked up and propagated in 2i+LIF. These iPSC lines were maintained under a 2i+LIF condition without feeders. The ESC and iPSC lines in this study have not been tested for mycoplasma contamination.

**Formation of reconstituted ovaries.** PGCLCs were differentiated from ESCs or iPSCs as described previously<sup>20</sup>, purified by FACSaria II (BD Bioscience) and aggregated with E12.5 female gonadal somatic cells in a low-binding U-bottom 96-well plate (NUNC) for 2 days of culture in GK15 supplemented with 1  $\mu$ M retinoic acid. To strictly remove residual PGCs from dissociated gonadal cells, both SSEA1 and CD31 antibodies (Miltenyi Biotec) were used according to the manufacturer's instructions. 5,000 PGCLCs were cultured with 50,000 gonadal somatic cells to produce one reconstituted ovary.

**IVDi culture.** Reconstituted ovaries were placed on Transwell-COL membranes (Coaster) soaked in  $\alpha$ MEM-based IVDi medium:  $\alpha$ MEM supplemented with 2% FCS, 150  $\mu$ M ascorbic acid (Sigma), 1  $\times$  Glutamax, 1  $\times$  penicillin/streptomycin and 55  $\mu$ M 2-mercaptoethanol (Life Technologies). At 4 days of culture, the culture medium was changed to StemPro-34-based IVDi medium: StemPro-34 SFM (Life Technologies) supplemented with 10% FCS, 150  $\mu$ M ascorbic acid, 1  $\times$  Glutamax, 1  $\times$  penicillin/streptomycin and 55  $\mu$ M 2-mercaptoethanol. From 7 days to 10 days of culture, 500 nM ICI182780 was added to the StemPro-34-based IVDi medium. At 21 days of culture, individual 2FLs were manually dissociated using sharpened tungsten needles.

**Isolation of 2FLs.** Electrically sharpened tungsten needles were used for the isolation of individual 2FLs. Interstitial cells between 2FLs were carefully removed. The 2FLs were separated from the rOvary and placed at largely regular intervals.

**IVG culture.** The single 2FLs on the Transwell-COL membranes were soaked in IVG- $\alpha$ MEM medium:  $\alpha$ MEM supplemented with 5% FCS, 2% polyvinylpyrrolidone (Sigma), 150  $\mu$ M ascorbic acid, 1  $\times$  Glutamax, 1  $\times$  penicillin/streptomycin, 100  $\mu$ M 2-mercaptoethanol, 55  $\mu$ g ml<sup>-1</sup> sodium pyruvate (Nacalai Tesque), 0.1 IU ml<sup>-1</sup> follicle-stimulating hormone (Follistim; MSD), 15 ng ml<sup>-1</sup> BMP15 and 15 ng ml<sup>-1</sup> GDF9 (R&D Systems). At 2 days of culture, BMP15 and GDF9 were withdrawn from the medium and then follicles were incubated in 0.1% TypeIV Collagenase (MP Biomedicals). After washing with  $\alpha$ MEM supplemented with 5%

FCS several times, the follicles were cultured in IVG- $\alpha$ MEM without BMP15 and GDF9. At 11 days of culture, cumulus-oocyte complexes grown on the membrane were picked up by a fine glass capillary.

**IVM culture.** Cumulus-oocyte complexes were transferred to IVM medium:  $\alpha$ MEM supplemented with 5% FCS, 25  $\mu$ g ml<sup>-1</sup> sodium pyruvate, 1  $\times$  penicillin/streptomycin, 0.1 IU ml<sup>-1</sup> follicular-stimulating hormone, 4 ng ml<sup>-1</sup> EGF, and 1.2 IU ml<sup>-1</sup> hCG (gonadotropin; ASKA). At 16 h of culture, swollen cumulus cells were stripped from the oocytes by treating with hyaluronidase (Sigma), and then MII oocytes were determined by 1st polar body extrusion. Due to concerns of damages from a short-wavelength light (405 nm) for detection of SC, a small number of SC-negative oocytes (an estimated 5% of total oocytes) were not removed but subjected simultaneously to the following IVF experiments. For experiments that need to purify SC-positive oocytes, such as analyses of the diameters, imprinting gene loci, the chromosome number and transcriptome, SC-positive oocytes were carefully collected under a fluorescent microscopy with the short-wavelength light. The contamination rate of the endogenous ICR MII oocytes resulted from the selection in these experiments.

**IVF and embryo transfer.** MII oocytes were subjected to IVF as described previously and then fertilized eggs were cultured in HTF medium (ARK Resource). 2-cell embryos were transferred into the oviduct of 0.5 days post coitum (d.p.c.) pseudopregnant ICR females. Some of the 2-cell embryos were cultured in KSOM (ARK Resource) for 4 days until the blastocyst stage. Newborns were delivered from the mother by Caesarean section and nursed by surrogate ICR mothers.

**RNA-seq analysis.** A directional RNA-seq library was constructed as described previously<sup>23</sup>. Briefly, poly(A)<sup>+</sup> RNAs were purified from 100–15,000 cells or oocytes (ESCs, 10,000–13,200; PGCLC-d6, 10,000; PGCLC-agg3, 5,300–15,000; vitro-2nd and vivo-2nd: 337–12; vitro-MII and vivo-MII: 99–134), using a Dynabeads mRNA DIRECT Micro Kit (Invitrogen). For vitro-2nd and vitro-MII, SC-positive oocytes were collected under a fluorescence microscopy. Biologically triplicated samples were prepared in each stage. Purified RNAs were subjected to library construction using a NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB). cDNAs were enriched by 12-cycle PCR. Sequencing of the libraries was performed with HiSeq2000 (Illumina). Obtained reads were processed with the FASTX tool kit<sup>24</sup> to remove short (<20 bp) and low quality (quality score <20) reads, followed by trimming of the adaptor sequence. Processed reads were mapped to the mouse mm10 genome using TopHat2/Bowtie2<sup>25</sup>. Cuffdiff<sup>26</sup>, a program in the Cufflinks software, and QuasR<sup>27</sup>, a Bioconductor package, were used for the differential gene and repetitive sequence expression analyses. Hierarchical clustering and principal component analysis were performed with R, based on RefSeq gene expression levels.

**Generation of rESCs.** Derivation of rESCs was performed using 2i+LIF<sup>17</sup> medium. Briefly, fertilized eggs derived from *in-vitro*-generated oocytes were cultured in KSOM until the blastocyst stage. The blastocysts were placed on MMC-treated mouse embryonic fibroblasts and cultured in 2i+LIF. After 5 to 7 days of culture, colonies were picked up and passaged several times to establish rESCs. During establishment, rESCs were passaged on a culture plate coated with 300 ng ml<sup>-1</sup> laminin (BD Bioscience) to remove mouse embryonic fibroblasts.

**Chimaera analysis.** rESCs were injected into the wild-type blastocysts. The blastocysts were transferred into the uterus of 2.5 d.p.c. pseudopregnant females and the embryos were collected at 10 days after transplantation. The embryos and their gonads were observed under a fluorescent microscope (SZX16; Olympus). For analysis of the contribution of rESCs to various tissues, the neural tube, lung, heart, liver, intestine and tail were collected and subjected to PCR analyses.

**Genotyping.** Genomic DNAs isolated from the tails of offspring, rESC lines and fetal tissues were subjected to PCR amplification using the primers (Supplementary Table 2). For sexing, PCR reaction amplified a 253 bp fragment from the X chromosome and 355 bp and 399 bp fragments from the Y chromosome. PCR products were resolved through electrophoresis in agarose gels and illuminated by ethidium bromide (EtBr).

**COBRA analysis and bisulfite sequencing.** Genomic DNAs were isolated from more than 200 SC-positive MII oocytes generated *in vitro* and wild-type MII oocytes *in vivo*, and from tails of offspring from *in-vitro*-derived oocytes and wild-type mice. Bisulfite reactions were performed using a MethyLamp DNA Modification Kit (Epigentek). For COBRA analysis, the DMRs of *H19*, *Igf2r*, *Peg3* and *Snrpn* were amplified with specific primers (Supplementary Table 2), as described previously<sup>28,29</sup>. The amplified DNA was digested with the following restriction enzymes (NEB): PvuI-HF for *H19*, HpyCH4IV for *Igf2r* and *Peg3*, and AciI for *Snrpn*. The digested samples were resolved through electrophoresis in 2% or 3% agarose gels and illuminated by EtBr. The PCR products were subcloned into the pGEM-T Easy vector (Promega) and were sequenced. For bisulfite sequencing of genomic DNA from oocytes, PCR amplification of the DMRs of *H19* and *Igf2r* was carried out as described previously<sup>13</sup>. The PCR products were subcloned into

the pGEM-T Easy vector (Promega) and were sequenced. The sequences were analysed and aligned by QUMA software<sup>30</sup>.

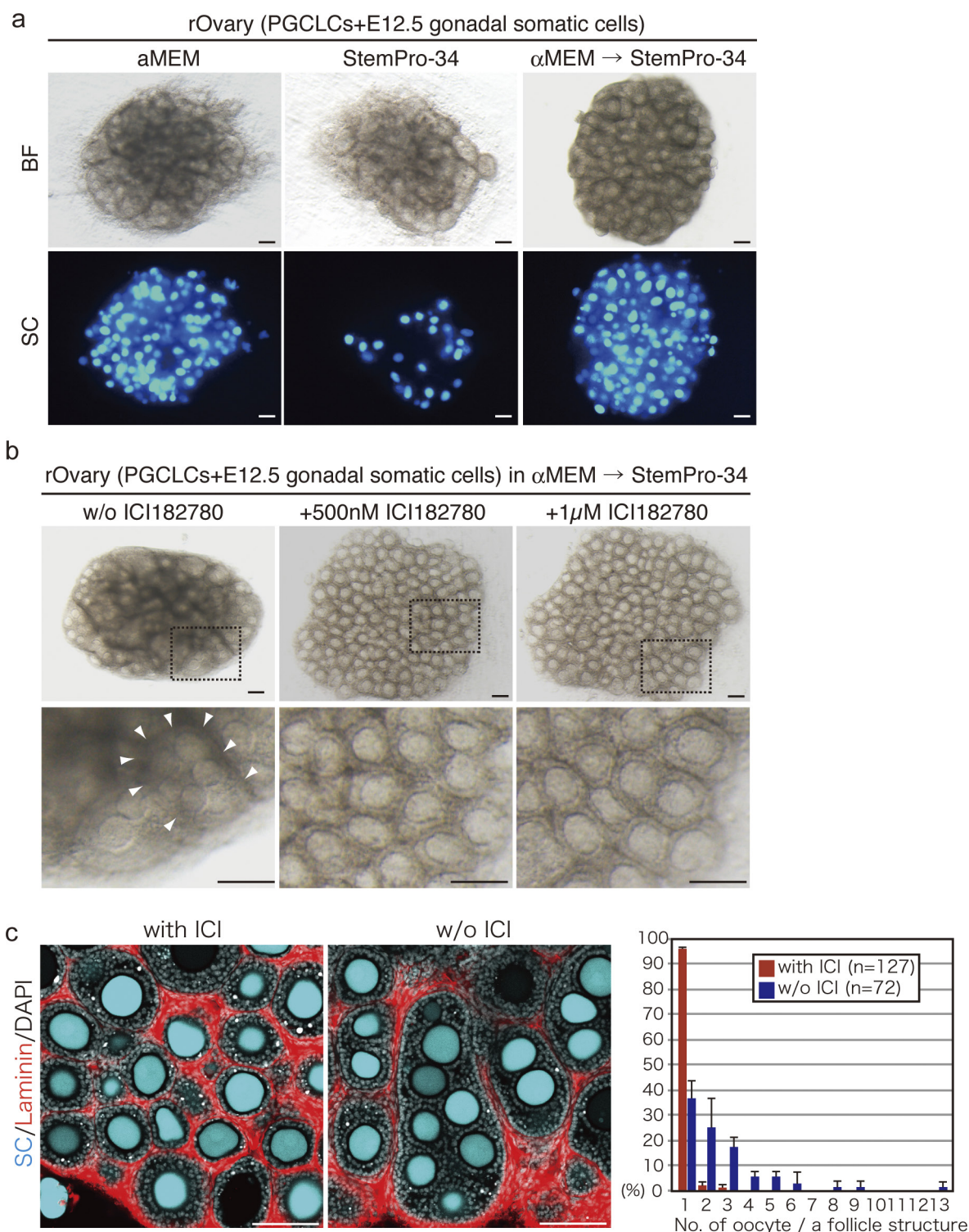
**Immunofluorescence analysis.** For whole-mount immunofluorescence analysis, rOvaries were fixed in 4% paraformaldehyde for 1 h at room temperature, washed three times in PBST (PBS containing 0.2% Triton X-100), soaked in blocking buffer (PBS containing 0.1% BSA and 0.3% Triton X-100) overnight at 4 °C, and then incubated with primary antibodies (anti-GFP rat monoclonal antibody (mAb) (Nakalai 04404-84) cross-reactive to BV and SC, and anti-Foxl2 goat polyclonal antibody (pAb) (Novus Biologicals NB100-1277) or anti-laminin rabbit pAb (Abcam ab11575)) in the blocking buffer overnight at 4 °C. After washing five times in PBST, the rOvaries were incubated with secondary antibodies (anti-rat IgG donkey pAb Alexa488 and anti-goat IgG donkey pAb Alexa568 or anti-rabbit IgG donkey pAb Alexa568) and DAPI in the blocking buffer overnight at 4 °C. The rOvaries were washed five times in PBST and mounted in Fluo-KEEPER antifade reagent (Nakalai).

Immunofluorescence of the chromosome in meiotic prophase I was performed as described previously<sup>31</sup> with slight modifications. Briefly, for the chromosome in meiotic prophase I, rOvaries were dissociated by incubation in CTK (0.1 mg ml<sup>-1</sup> collagenase IV, 0.25% Trypsin, 20% KSR and 1 mM CaCl<sub>2</sub> in PBS) for 30 min at 37 °C, followed by Accutase (Nakalai) for 5 min at 37 °C. Dissociated single cells were suspended in hypotonic buffer and placed on glass slides. Slides were washed three times in PBS, soaked in blocking buffer (PBS containing 10% FBS) for 1 h at room temperature and then incubated with primary antibodies (anti-SCP1 rabbit pAb (Novus Biologicals NB300-ss9) and anti-SCP3 mouse mAb (Abcam ab97672)) overnight at 4 °C. After washing three times in PBS containing 0.05% Tween20, the slides were incubated with secondary antibodies (anti-rabbit IgG goat pAb Alexa568 and anti-mouse IgG donkey pAb Alexa647), washed three times in PBS containing 0.05% Tween20, and then incubated with anti-γH2AXSer139 mouse mAb directly conjugated with FITC (Millipore 16-202A) and DAPI. The slides were washed three times in PBS containing 0.05% Tween20 and mounted in Fluo-KEEPER antifade reagent (Nakalai). Immunofluorescence of the chromosome in MII oocytes was performed as described previously<sup>32</sup>. The primary antibodies

used were anti-centromere human pAb (Antibodies Incorporated 15-235-001) and anti-H3K9me3 rabbit pAb (Millipore 07-442). The secondary antibodies used were anti-human IgG donkey pAb Alexa488 and anti-rabbit IgG goat pAb Alexa568. All immunofluorescence samples were analysed by a confocal microscope (Zeiss LSM700).

20. Hayashi, K. & Saitou, M. Generation of eggs from mouse embryonic stem cells and induced pluripotent stem cells. *Nat. Protocols* **8**, 1513–1524 (2013).
21. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
22. Nagamatsu, G. *et al.* A germ cell-specific gene, *Prmt5*, works in somatic cell reprogramming. *J. Biol. Chem.* **286**, 10641–10648 (2011).
23. Hamazaki, N., Uesaka, M., Nakashima, K., Agata, K. & Imamura, T. Gene activation-associated long noncoding RNAs function in mouse preimplantation development. *Development* **142**, 910–920 (2015).
24. Patel, R. K. & Jain, M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* **7**, e30619 (2012).
25. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
26. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protocols* **7**, 562–578 (2012).
27. Gaidatzis, D., Lerch, A., Hahne, F. & Stadler, M. B. QuasR: quantification and annotation of short reads in R. *Bioinformatics* **31**, 1130–1132 (2015).
28. Li, X. *et al.* A maternal-zygotic effect gene, *Zfp57*, maintains both maternal and paternal imprints. *Dev. Cell* **15**, 547–557 (2008).
29. Lee, J. *et al.* Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development* **129**, 1807–1817 (2002).
30. Kumaki, Y., Oda, M. & Okano, M. QUMA: quantification tool for methylation analysis. *Nucleic Acids Res.* **36**, W170–W175 (2008).
31. Takada, Y. *et al.* HP1γ links histone methylation marks to meiotic synapsis in mice. *Development* **138**, 4207–4217 (2011).
32. Hodges, C. A. & Hunt, P. A. Simultaneous analysis of chromosomes and chromosome-associated proteins in mammalian oocytes and embryos. *Chromosoma* **111**, 165–169 (2002).





# Extended Data Figure 1 | Evaluation of the basal culture conditions.

**a**, Culture of rOvaries. Representative images of rOvaries at 3 weeks of culture in the medium indicated on the top are shown. Note that a number of SC-positive oocytes were formed with firm structure of rOvary under the combined culture condition. BF, bright field; SC, *stella-ECFP*.

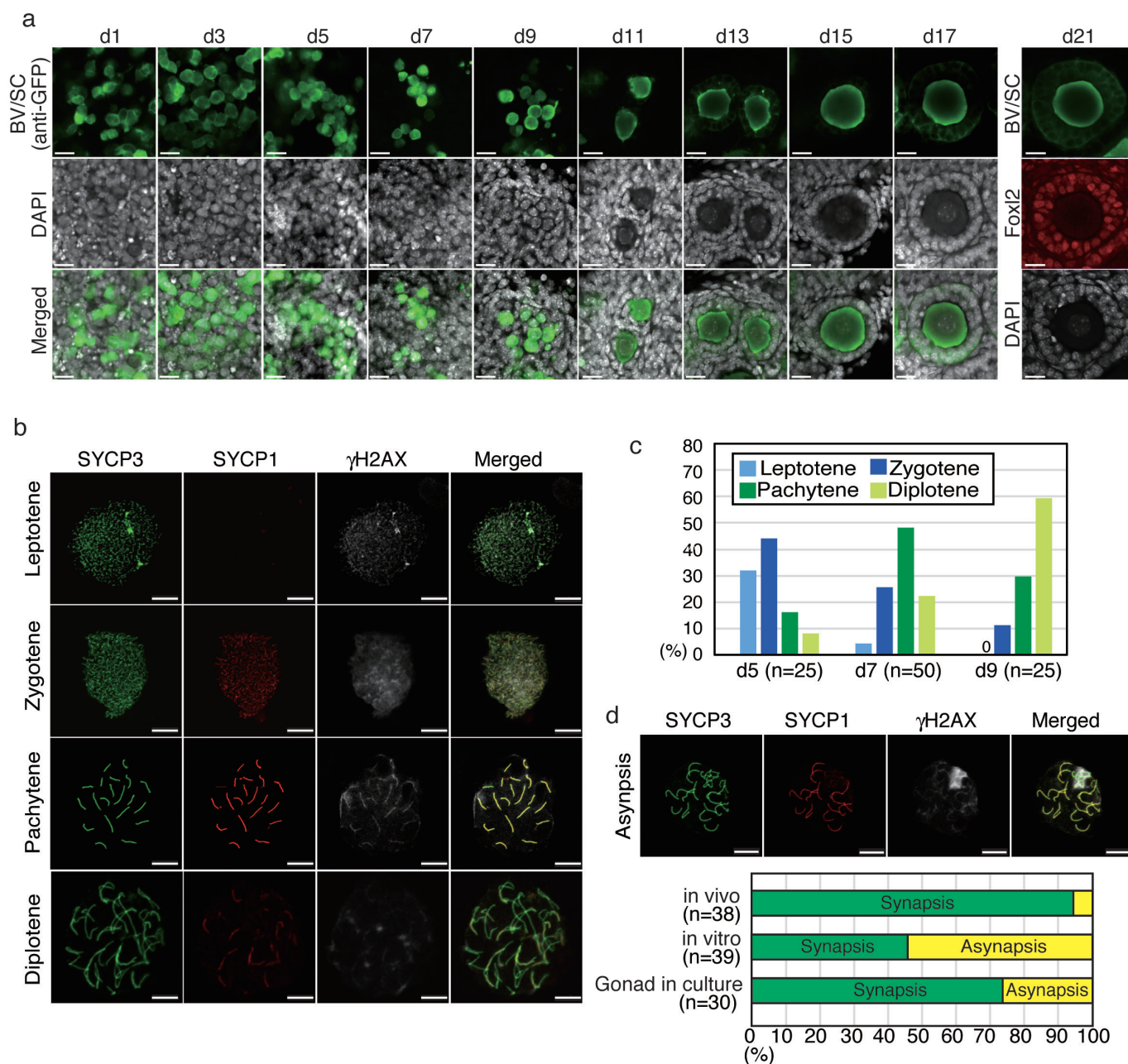
**b**, Culture of rOvaries with ICI182780, an oestrogen inhibitor.

Representative images of rOvaries under the culture condition indicated on the top are shown. The images on the bottom of the figure are high-magnification images of the region in the dashed box. Note that multiple oocyte follicles (white arrowheads) were frequently observed in the

culture without ICI182780 but were rarely observed in the culture with the inhibitor. **c**, Immunofluorescence analysis of the follicle structure.

Images are rOvaries at 3 weeks of culture with or without ICI182780.

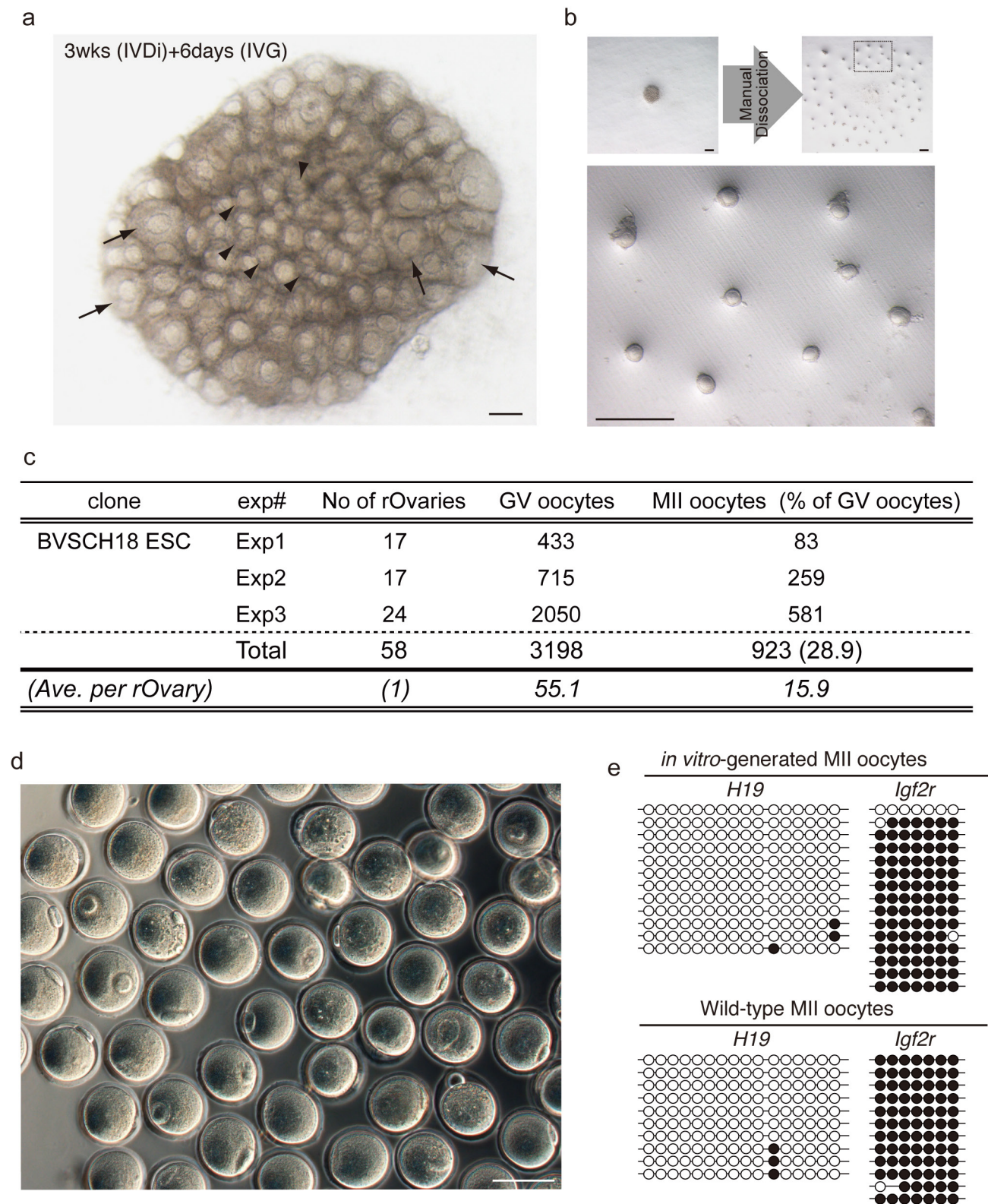
Immunostaining with laminin, a component of the basement membrane, was used to define the follicle structure that encloses oocyte(s). The bar graph on the right shows the percentage of follicles with single or multiple oocytes. The mean values  $\pm$  s.d. from 3 biological replicates are shown. The total numbers of follicles counted are shown in the box. Scale bars, 100  $\mu$ m.



**Extended Data Figure 2 | Meiotic progression and establishment of genomic imprints in culture.** **a**, PGC-to-oocyte differentiation during IVDi. PGCLCs and oocytes were visualized by anti-GFP antibody (green) cross-reactive to BV and SC. At 21 days of culture, cells surrounding the oocyte expressed Foxl2 (red). Scale bars, 20  $\mu$ m. **b**, Representative images of each stage of meiotic prophase I in culture. Meiotic cells from the culture were spread and stained with the antibodies indicated. The colours were converted by ZEN software (Zeiss). **c**, Meiotic progression in IVDi culture. Percentages of each stage in total meiotic cells are shown.

The meiotic cells in the rOvaries were obtained from 3 independent experiments. **d**, Asynapsis of meiotic chromosome at the pachytene stage in culture. Representative immunofluorescence images of asynapsis at the pachytene stage are shown. Note that  $\gamma$ H2AX is densely accumulated in the meiotic chromosome. The graph below the images shows the percentage of asynapsis in cells at the pachytene stages in three E17.5 female gonads, four rOvaries or three E12.5 gonads cultured for 5 days, from 3 independent experiments. Scale bars in **b** and **d**, 10  $\mu$ m.

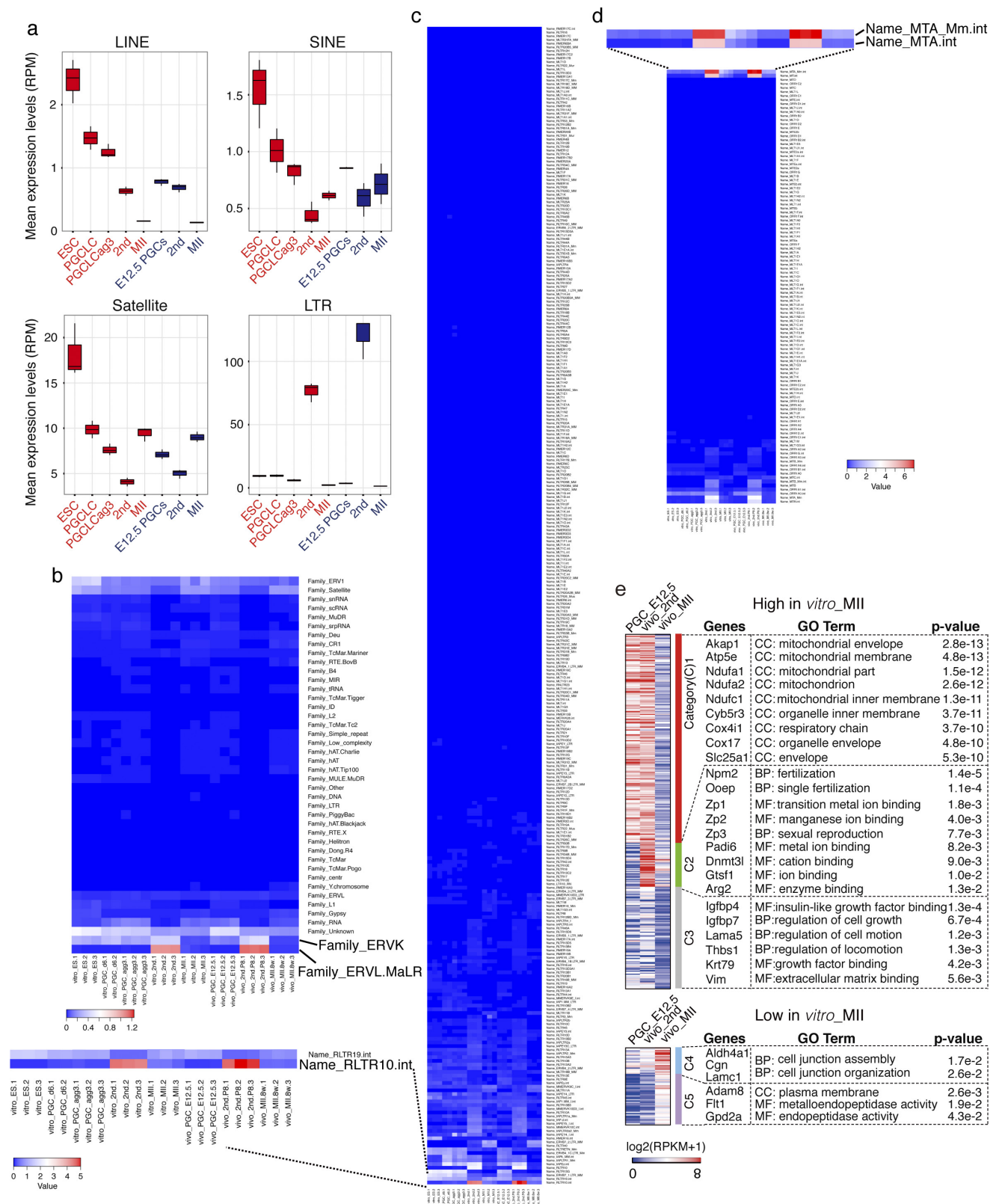




**Extended Data Figure 3 | Isolation of individual 2FLs for IVG culture and MII oocytes from IVM culture.** **a**, Attenuated maturation in an intact rOvary. Cell proliferation in the granulosa cell layer is observed only in 2FLs located at the edge of the rOvary (arrows). 2FLs in the middle seldom show cell proliferation in the granulosa cell layer (arrowheads). **b**, Representative images of isolation of individual 2FLs. A rOvary (left, upper) was mechanically dissociated (right upper). The high-magnification image of the boxed region shows individual 2FLs

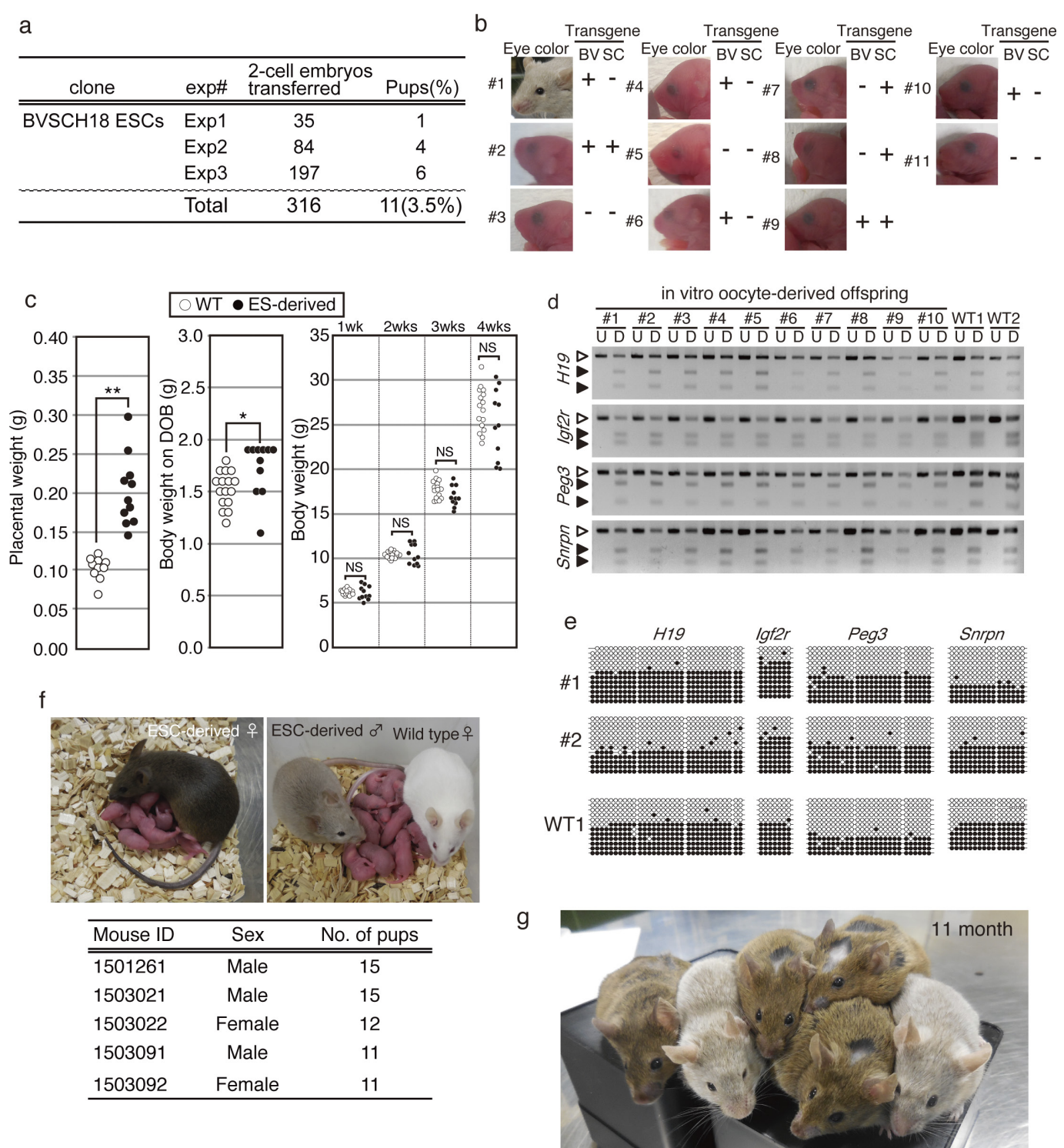
after dissociation. **c**, The number of germinal vesicle and MII oocytes derived from BVSCH18 ESCs. The number of oocytes obtained from each experiment and the total number from all experiments were shown. **d**, A large image of the MII oocytes shown in Fig. 1d. **e**, Bisulfite sequence analysis of differentially methylated regions (DMRs) of the imprinted genes (*H19* and *Igf2r*) in MII oocytes generated *in vitro* and *in vivo*. White and black circles represent unmethylated and methylated CpG sequences, respectively. Scale bars, 100  $\mu$ m (**a**, **d**), 500  $\mu$ m (**b**).





**Extended Data Figure 4 | Expression of a specific class of LTR.** **a**, Box plots showing the expression of repetitive elements. Reads per million mapped reads (RPMs) of each product are shown. Note that LTR products increased sharply in oocytes at the secondary follicle stage. **b**, **c**, **d**, Heat map of the normalized expression profile of the LTR family (**b**), ERVK

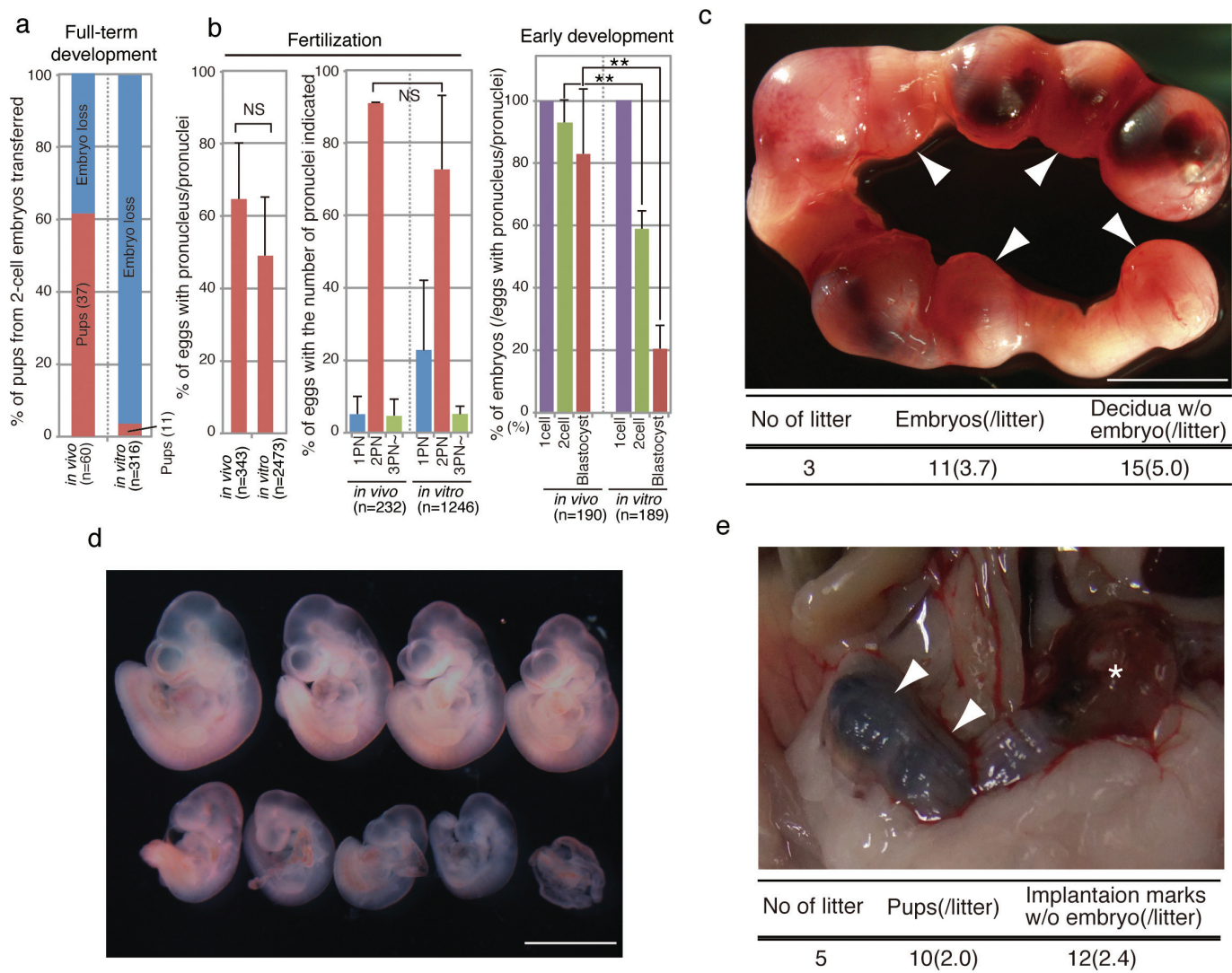
(**c**) and ERVL-MaLR (**d**). **e**, The gene ontology enrichments of genes categorized by their gene expression dynamics. All genes listed are shown in Supplementary Table 1. All expression data are based on the mean values from 3 biological replicates except for PGCLC(d6) that is from 2 biological replicates.



**Extended Data Figure 5 | Growth and fertility of offspring from *in-vitro*-generated oocytes.** **a**, The number of 2-cell embryos and pups derived from BVSCH18 ESCs. **b**, Summary of eye colour and transgene of the pups (or adult mouse) derived from *in-vitro*-generated oocytes from BVSCH18 ESCs. **c**, Weights of placentas (left), newborn pups (middle) and development of the body weights (right) of offspring from the MII oocytes from BVSCH18 (closed circles) and the genetically matched wild-type mice (129X1/svjC57BL/6F1 × ICR) (closed circles) in 3 independent experiments (*t*-test, \*\**P* < 0.01, \**P* < 0.05). **d**, Combined bisulfite restriction analysis (COBRA) of DMRs of the imprinted genes (*H19*, *Igf2r*, *Peg3* and *Snrpn*) in the 10 mice (1–10) derived from *in-vitro*-

generated MII oocytes and the two wild-type mice (WT1 and WT2). PCR products were either digested (D) or undigested (U) with the respective enzyme. The digested and undigested fragments are indicated by black and white triangles, respectively. **e**, Bisulfite sequence analysis of DMRs of the imprinted genes in the two mice (1 and 2) from *in-vitro*-generated oocytes and the wild-type mouse (WT1). White and black circles represent unmethylated and methylated CpG sequences, respectively. **f**, A female (left) and a male (right) mouse from *in-vitro*-generated oocytes with full fertility. The table below the images shows the number of pups in a litter. **g**, The 6 mice derived from *in-vitro*-generated MII oocytes at 11 months after birth. For gel source data see Supplementary Fig. 1.

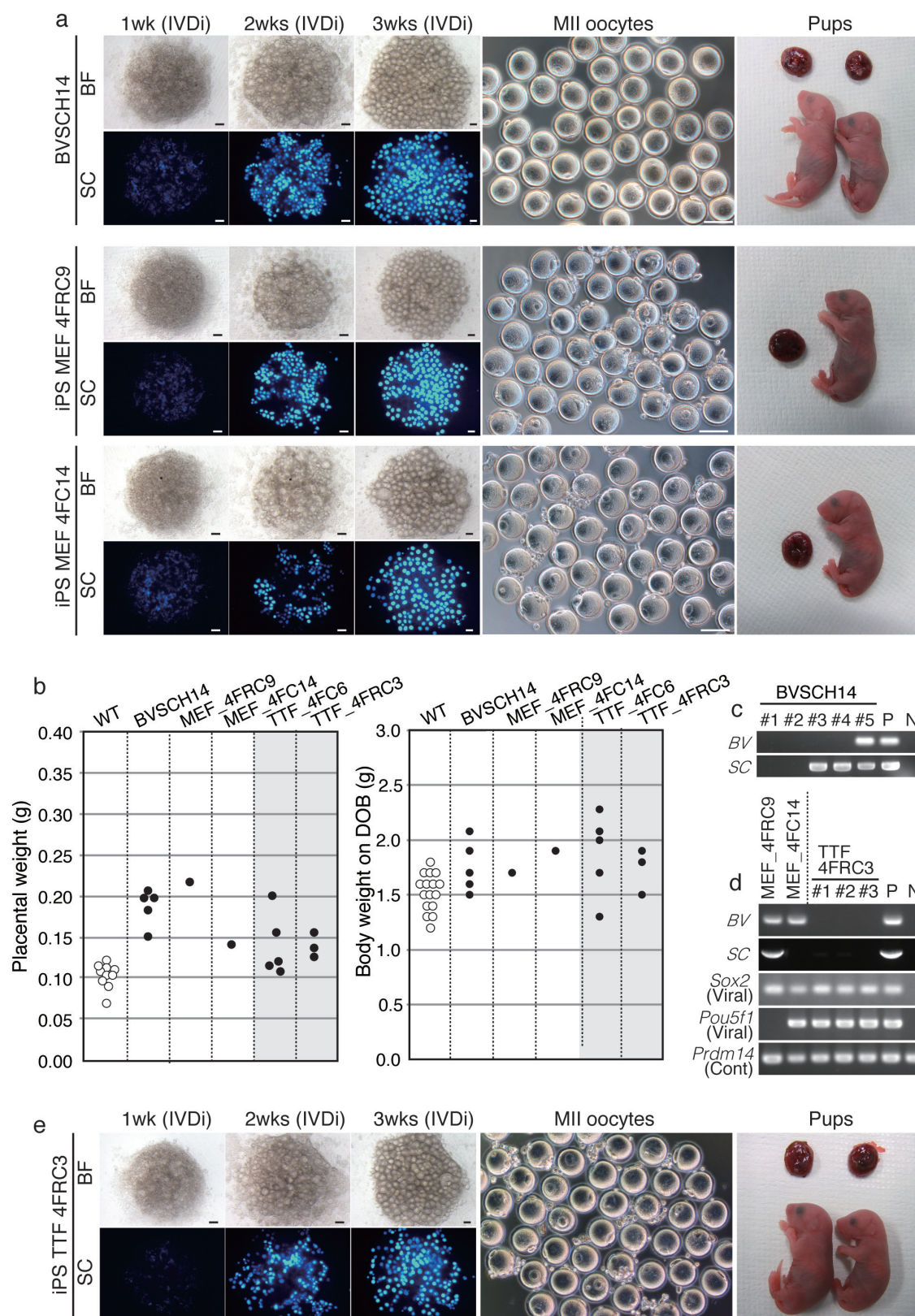




**Extended Data Figure 6 | Developmental potential of *in-vitro*-generated oocytes.** **a**, Percentages of pups from 2-cell embryos transferred. The numbers of pups obtained from 3 independent embryo transfer experiments are shown. **b**, Fertilization and preimplantation development. Eggs with pronucleus/pronuclei (PN) (left), the number of PN (middle) and early embryonic development in culture (right). The mean values  $\pm$  s.d. from 5 biological replicates are shown. The total numbers of oocytes/embryos are shown in each graph. **c**, A uterus at 10 days after transfer of embryos from ESC-derived oocytes. White arrowheads indicate degenerating conceptuses that had decidua without any apparent embryo, and therefore embryo loss at early gestation. The table below the image summarizes the numbers of remaining embryos and decidua

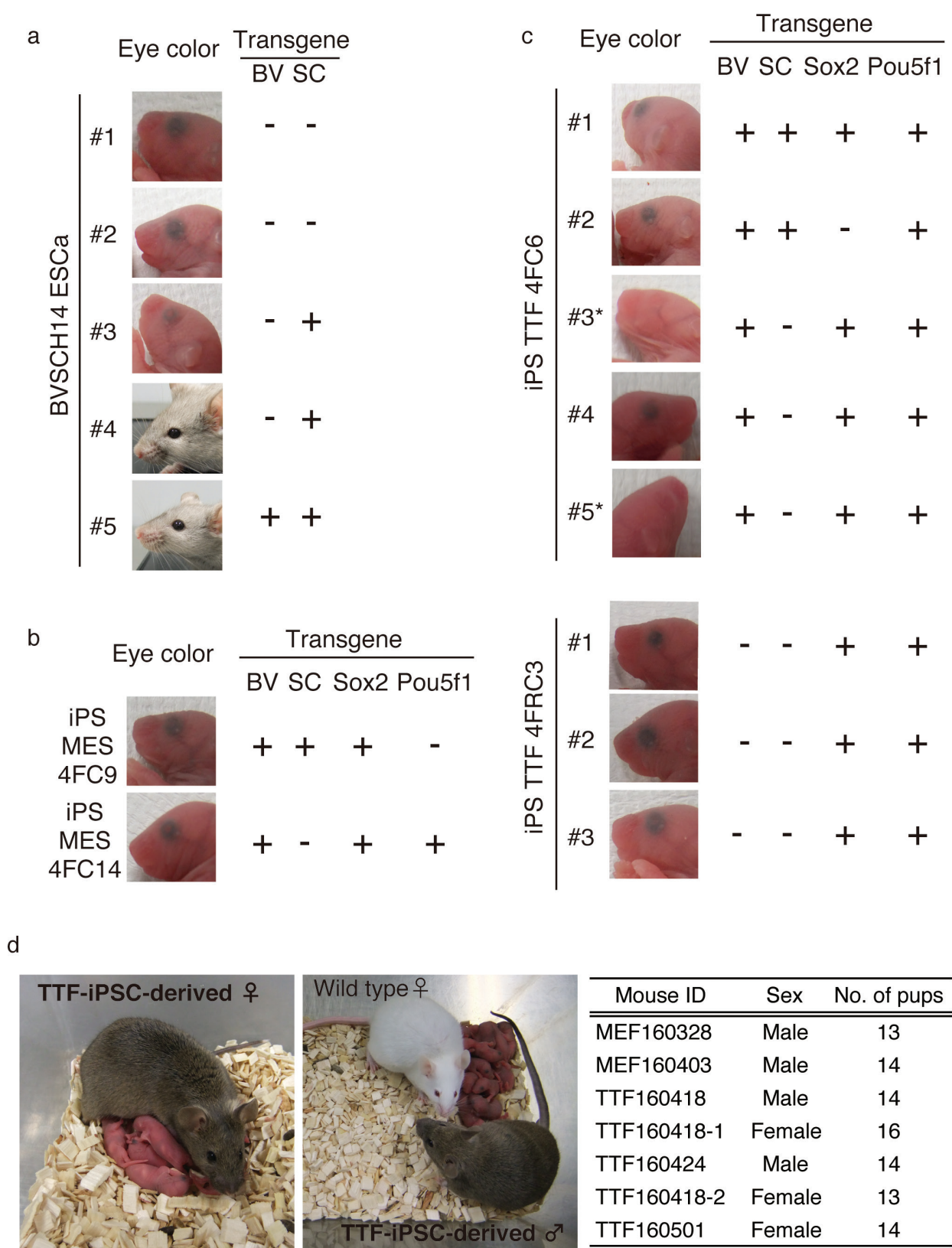
without embryo observed at 10 days of gestation. On average, 3.7 embryos remained in the uterus. Scale bar, 5 mm. **d**, Embryos collected from a day 10 uterus after transfer of embryos from ESC-derived oocytes. The developmental stage of the embryos was varied. Scale bar, 2 mm. **e**, Implantation marks in the uterus on Caesarean section. White arrowheads indicate the implantation marks without embryos and the asterisk shows the uterus wall after delivering newborn pups by Caesarean section. The table below the image summarizes the numbers of pups and implantation marks without embryos on Caesarean section. On average, 2 pups were obtained per uterus, indicating there was some amount of embryo loss during late gestation.





**Extended Data Figure 7 | Oogenesis *in vitro* using other ESC and iPSC lines.** **a**, Representative images of a rOvary in IVDi, MII oocytes and pups obtained from BVSCH14 ESCs and MEF-derived iPSC lines (MEF4FRC9 and MEF4FC14). **b**, Weights of placentas (left) and newborns (right) from the MII oocytes from ESC/iPSC lines indicated on the top. The values are from 2 independent experiments (see also Extended Data Table 1). The values of the control are same as those shown in Extended Data

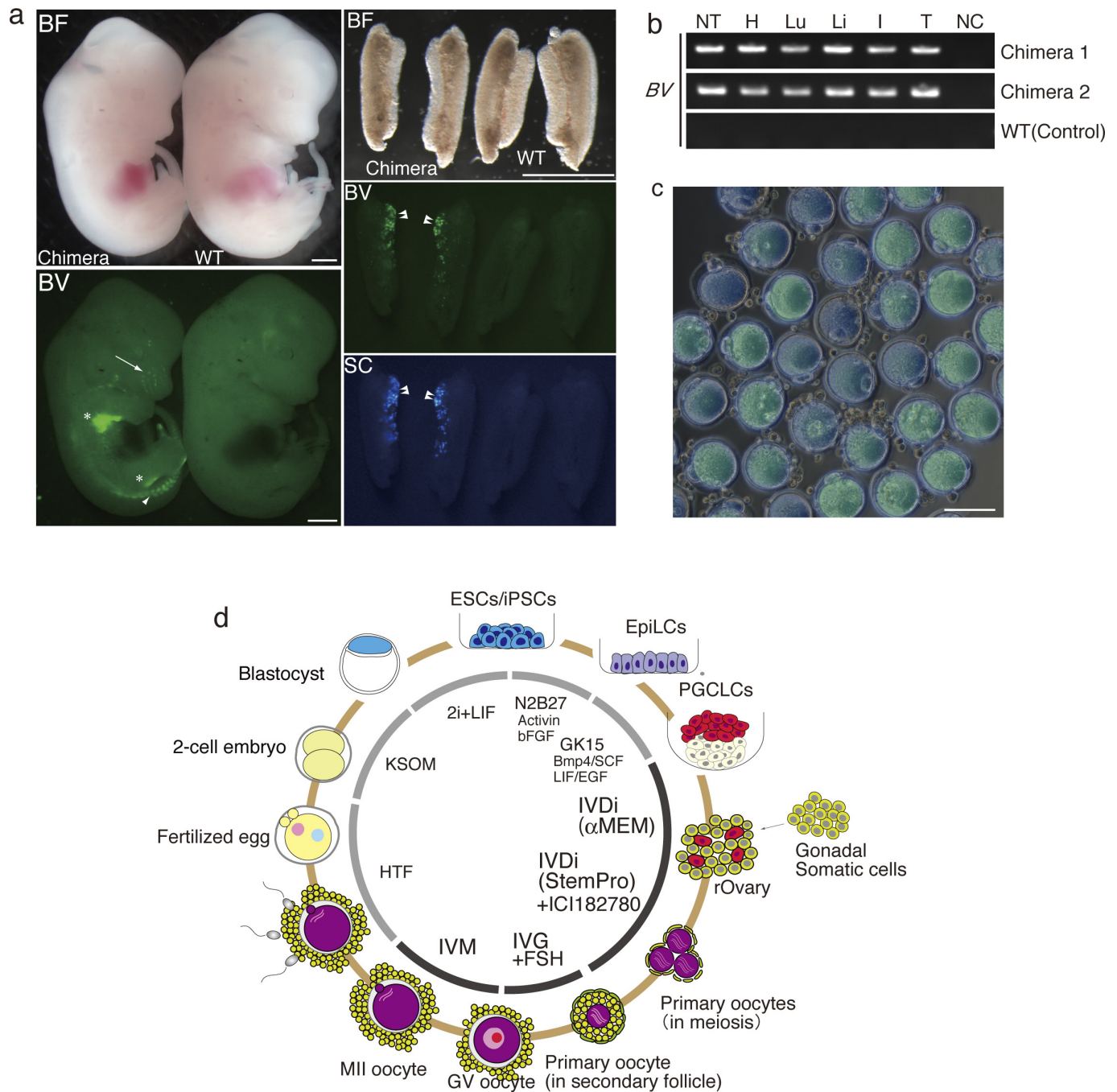
**Fig. 5c, d**, Genotyping of the pups from BVSCH14 ESCs (**c**) and iPSCs (**d**). PCR products from each gene indicated are shown. P, a positive control; N, a negative control. Details of the positive control and negative control are described in Fig. 3 legend. **e**, Representative images of a rOvary in IVDi using TTF-derived iPSCs (TTF4FRC3), MII oocytes and pups obtained from TTF4FRC3. Scale bars, 100  $\mu$ m. For gel source data see Supplementary Fig. 1.



**Extended Data Figure 8 | Summary of eye colour and transgene of the pups derived from *in-vitro*-generated oocytes.** **a–c,** A compiled list of eye colour and transgene of the pups derived from *in-vitro*-generated oocytes from BVSCH14 ESCs (**a**), MEF-derived iPSCs (MEF4FRC9 and MEF4FC14) (**b**), and TTF-derived iPSCs (TTF4FC6 and TTF4FRC3) (**c**). 3\* and 5\* pups from TTF4FC6 had albino eyes. This is because MEF-iPSCs and TTF-iPSCs (129X1/svj (chinchilla) × C57BL/6) used in this

study have Cc or Cc<sup>ch</sup> genotype that effects eye pigmentation (see also Method). As ICR strain has cc genotype, half of the offspring would have albino eyes. Note that these pups had BV and the retroviral transgenes. **d,** Fertility of adult mice derived from *in-vitro*-generated oocytes from iPSCs. Both a female (left) and a male (middle) from TTF-derived iPSCs were fertile. The table (right) shows the number of pups in a litter.





**Extended Data Figure 9 | Chimaera analysis of rESCs.** **a**, Bright field (BF) and fluorescent image (BV or SC) of a representative chimera and wild-type E12.5 embryos (left images), and their gonads with mesonephros (right images). BV expression was detectable in the sensory vibrissae (arrow), the mesenchyme of the forelimb and the hind limb (asterisks) and the myotome (arrowheads), as reported previously<sup>10</sup>. BV and SC expressions were detectable in PGCs in the gonads (double

arrowheads). Scale bars, 1 mm. **b**, BV transgene in chimera embryos. PCR products from each tissue (NT, neural tube; H, heart; Lu, lung; Li, liver; I, intestine; T, tail; NC, a negative control from the genomic DNA of a wild-type embryo) are shown. **c**, MII oocytes from rESCs. The image in Fig. 4d merged with SC is shown. Scale bars, 100  $\mu$ m. **d**, A schematic drawing showing the *in vitro* reconstitution of the entire female germ line. For gel source data see Supplementary Fig. 1.



**Extended Data Table 1 | Summary of oogenesis *in vitro* using BVSC14 ESCs and iPSCs derived from MEFs and TTFs**

clone	exp#	agg	GV oocytes (/rOvary)	MII oocyte (/rOvary)	Egg with PN (1PN, 2PN, 3PN)	2cell embryos (% of Egg with PN)	2cell embryos transferred <sup>a</sup>	pups (% of 2cell embryos transferred)
BVSC14 ESCs	Exp1	16	1616	683	438 (28, 385, 25)	296	296	3
	Exp2	29	2465	802	400 (65, 310, 25)	267	133	2
	Total	45	4081 (90.7)	1485 (33.0)	838 (93, 695, 50)	563 (67.2)	429	5 (1.2)
iPS MEF_4FRC9	Exp1	35	2939	841	410 (51, 320, 39)	125	125	0
	Exp2	31	2620	710	381 (86, 278, 17)	241	241	1
	Total	66	5559 (84.2)	1551 (23.5)	791 (137, 598, 56)	366 (46.3)	366	1 (0.3)
iPS MEF_4FC14	Exp1	33	2691	727	404 (41, 330, 33)	104	94	0
	Exp2	31	2397	902	319 (102, 210, 7)	200	200	1
	Total	64	5088 (79.5)	1629 (25.5)	723 (143, 540, 40)	304 (42.0)	294	1 (0.3)
iPS TTF_4FC6	Exp1	32	3185	658	576 (71, 465, 40)	317	237	3
	Exp2	31	3123	1042	540 (124, 364, 52)	335	335	2
	Total	63	6308 (100.1)	1700 (27.0)	1116 (195, 829, 92)	652 (58.4)	572	5 (0.9)
iPS TTF_4FRC3	Exp1	30	3765	1138	877 (176, 606, 95)	335	335	1
	Exp2	31	N.C. <sup>b</sup>	1210	711 (166, 501, 44)	441	441	2
	Total	61	3765 (125.5)	2348 (38.5)	1588 (342, 1107, 139)	776 (48.9)	776	3 (0.4)

<sup>a</sup>As they were used for other experiments, not all 2-cell embryos were transferred into the surrogate mothers.

<sup>b</sup>Not counted.

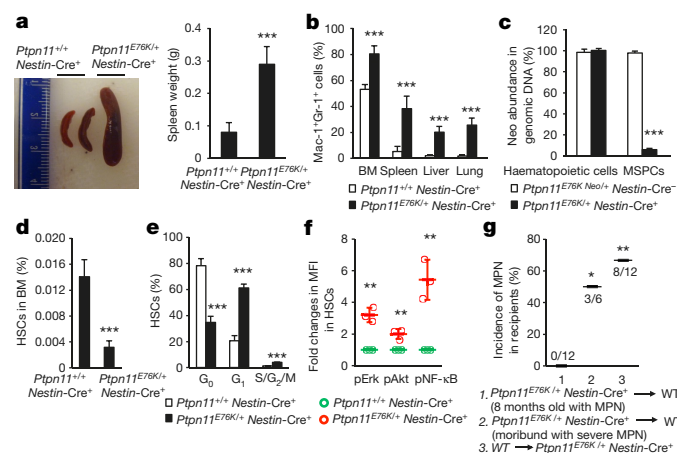
# Leukaemogenic effects of *Ptpn11* activating mutations in the stem cell microenvironment

Lei Dong<sup>1\*</sup>, Wen-Mei Yu<sup>1\*</sup>, Hong Zheng<sup>1\*</sup>, Mignon L. Loh<sup>2</sup>, Silvia T. Bunting<sup>3</sup>, Melinda Pauly<sup>1</sup>, Gang Huang<sup>4</sup>, Muxiang Zhou<sup>1</sup>, Hal E. Broxmeyer<sup>5</sup>, David T. Scadden<sup>6</sup> & Cheng-Kui Qu<sup>1</sup>

Germline activating mutations of the protein tyrosine phosphatase SHP2 (encoded by *PTPN11*), a positive regulator of the RAS signalling pathway<sup>1</sup>, are found in 50% of patients with Noonan syndrome<sup>2</sup>. These patients have an increased risk of developing leukaemia<sup>3</sup>, especially juvenile myelomonocytic leukaemia (JMML), a childhood myeloproliferative neoplasm (MPN). Previous studies have demonstrated that mutations in *Ptpn11* induce a JMML-like MPN through cell-autonomous mechanisms that are dependent on Shp2 catalytic activity<sup>4–7</sup>. However, the effect of these mutations in the bone marrow microenvironment remains unclear. Here we report that *Ptpn11* activating mutations in the mouse bone marrow microenvironment promote the development and progression of MPN through profound detrimental effects on haematopoietic stem cells (HSCs). *Ptpn11* mutations in mesenchymal stem/progenitor cells and osteoprogenitors, but not in differentiated osteoblasts or endothelial cells, cause excessive production of the CC chemokine CCL3 (also known as MIP-1 $\alpha$ ), which recruits monocytes to the area in which HSCs also reside. Consequently, HSCs are hyperactivated by interleukin-1 $\beta$  and possibly other proinflammatory cytokines produced by monocytes, leading to exacerbated MPN and to donor-cell-derived MPN following stem cell transplantation. Remarkably, administration of CCL3 receptor antagonists effectively reverses MPN development induced by the *Ptpn11*-mutated bone marrow microenvironment. This study reveals the critical contribution of *Ptpn11* mutations in the bone marrow microenvironment to leukaemogenesis and identifies CCL3 as a potential therapeutic target for controlling leukaemic progression in Noonan syndrome and for improving stem cell transplantation therapy in Noonan syndrome-associated leukaemias.

In our recent study investigating the potential effects of *Ptpn11* activating mutations in neural cells, we used the *Ptpn11*<sup>E76K</sup> mutation as a model and generated *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice with *Ptpn11*<sup>E76K</sup> mutation conditional knock-in mice (*Ptpn11*<sup>E76K-neo/+</sup>)<sup>5</sup> and *Nestin-Cre*<sup>+</sup> mice. We inadvertently found that *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice developed a myeloid malignancy resembling MPN at the age of 7 months or older as evidenced by splenomegaly, and significantly increased numbers of myeloid cells in the peripheral blood and myeloid progenitors in the bone marrow (BM) (Fig. 1a, Extended Data Fig. 1a, b). Histopathological examination revealed hyperproliferation of myeloid cells in the BM and spleen (Extended Data Fig. 1c). Myeloid cells (Mac-1<sup>+</sup>Gr-1<sup>+</sup>) (Fig. 1b) and inflammatory monocytes (CD115<sup>+</sup>Gr-1<sup>+</sup>) (Extended Data Fig. 1d) were significantly increased in these tissues. Moreover, extensive myeloid cell infiltration in the liver and lung was detected (Fig. 1b, Extended Data Fig. 1c). The *loxP*-flanked neo cassette with a stop codon, which inactivated the targeted *Ptpn11*<sup>E76K-neo</sup> allele<sup>5</sup>, was intact in the MPN cells of these

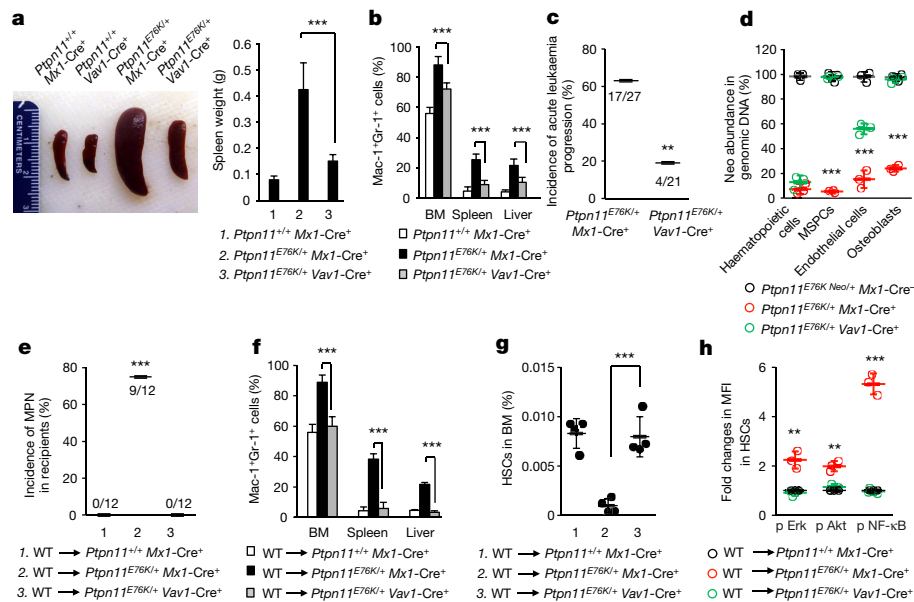
mice (Fig. 1c), indicating that the myeloid malignancy was not caused by the *Ptpn11* mutation in haematopoietic cells. Previous studies have shown that Nestin is also expressed in BM mesenchymal stem/progenitor cells (MSPCs) in addition to neural cells, and that perivascular Nestin<sup>+</sup> MSPCs constitute unique sinusoidal vascular and arteriolar HSC niches<sup>8,9</sup>. We therefore examined targeted *Ptpn11* alleles in BM-derived MSPCs and found that the inhibitory neo cassette was deleted in approximately 95% of these cells (Fig. 1c). Interestingly, the frequency and absolute numbers of primitive haematopoietic progenitors and stem cells in the BM were markedly decreased in



**Figure 1 | *Ptpn11*<sup>E76K/+</sup> mutation in Nestin<sup>+</sup> MSPCs aberrantly activates neighbouring wild-type HSCs, inducing MPN in *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice.** 7–12-month-old *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> and *Ptpn11*<sup>E76K-neo/+</sup>*Nestin-Cre*<sup>+</sup> mice were analysed. **a**, Spleen weights were determined ( $n = 17$  mice per group). **b**, Cells isolated from BM, spleens, livers and lungs were assayed for Mac-1<sup>+</sup>Gr-1<sup>+</sup> myeloid cells by FACS ( $n = 12$  mice per group). **c**, Genomic DNA isolated from BM haematopoietic cells and BM-derived MSPCs was assayed for the abundance of the neo cassette by qPCR ( $n = 5$  mice per group). **d–f**, BM cells were assayed by multiparameter FACS to determine the pool size ( $n = 8$  mice per group) (**d**), cell cycle distribution ( $n = 6$  mice per group) (**e**), and intracellular signalling activities ( $n = 3$  mice per group) (**f**) of HSCs (Lin<sup>+</sup>Sca-1<sup>+</sup>c-Kit<sup>+</sup>CD150<sup>+</sup>CD48<sup>+</sup>Flk2<sup>+</sup>). **g**, BM cells collected from 8-month-old *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice (CD45.2<sup>+</sup>) with MPN or moribund *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice (12 months old) with severe MPN were transplanted into lethally irradiated wild-type mice (WT) Boyl mice (CD45.1<sup>+</sup>). In addition, BM cells collected from Boyl mice were transplanted into 6-month-old *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice. Recipients were monitored for MPN development for 6–8 months. Data shown in **a–f** are mean  $\pm$  s.d. of all mice examined; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . Source Data for this figure are available online.

<sup>1</sup>Department of Pediatrics, Division of Hematology/Oncology, Aflac Cancer and Blood Disorders Center, Children's Healthcare of Atlanta, Emory University School of Medicine, Atlanta, Georgia 30322, USA. <sup>2</sup>Department of Pediatrics, Division of Pediatric Hematology-Oncology, University of California at San Francisco, San Francisco, California 94122, USA. <sup>3</sup>Department of Pathology, Children's Healthcare of Atlanta, Emory University, Atlanta, Georgia 30322, USA. <sup>4</sup>Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital, University of Cincinnati, Cincinnati, Ohio 45229, USA. <sup>5</sup>Department of Microbiology and Immunology, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA. <sup>6</sup>Center for Regenerative Medicine and MGH Cancer Center, Massachusetts General Hospital, Department of Stem Cell and Regenerative Biology and Harvard Stem Cell Institute, Harvard University, Boston, Massachusetts 02114, USA.

\*These authors contributed equally to this work.



**Figure 2 | MPN that developed in *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> mice is less severe and less progressive than MPN developed in *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> mice.** *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>*Mx1-Cre*<sup>+</sup> mice (8 weeks after pI-pC administration), along with 16-week-old *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>*Vav1-Cre*<sup>+</sup> mice were killed. **a**, Spleen weights were determined ( $n = 8$  mice per group). **b**, Cells isolated from BM, spleens and livers were assayed for Mac-1<sup>+</sup>Gr-1<sup>+</sup> myeloid cells ( $n = 8$  mice per group). **c**, *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> and *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> mice administered with pI-pC were monitored for 12 months for acute leukaemia progression. **d**, Haematopoietic cells (CD45<sup>+</sup>), MSPCs (Sca-1<sup>+</sup>CD140α<sup>+</sup>CD45<sup>-</sup>Ter-119<sup>-</sup>CD31<sup>-</sup>), endothelial cells (CD45<sup>-</sup>Ter-119<sup>-</sup>CD31<sup>+</sup>), and osteoblasts (Sca-1<sup>-</sup>CD140α<sup>+</sup>CD45<sup>-</sup>Ter-119<sup>-</sup>CD31<sup>-</sup>) were sorted from the BM.

*Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice, whereas these cells in the spleen were increased (Fig. 1d, Extended Data Fig. 1e–g). The fact that the numbers of mature myeloid cells and myeloid progenitors increased whereas stem cells decreased implied aberrant activation and accelerated differentiation of HSCs in the BM. Indeed, the number of quiescent HSCs in the G<sub>0</sub> phase in *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice decreased by twofold, whereas that of HSCs in the G<sub>1</sub> or S/G<sub>2</sub>/M phases doubled (Fig. 1e). HSCs in these mice had reduced apoptosis (Extended Data Fig. 1h). Assessment of intracellular signalling activities demonstrated that Erk, Akt and NF-κB pathways were highly activated in the HSCs of *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice (Fig. 1f). The MPN developed in chronic-phase *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice was not transferable to wild-type transplants, but MPN cells from terminally ill mice reproduced the same disease in 50% of the recipients (Fig. 1g), possibly owing to the acquisition of unknown genetic mutations that conferred self-renewal capability to MPN cells. At a 6–8-month follow up, 8 of 12 lethally irradiated *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice that were transplanted with wild-type BM cells developed donor-cell-derived MPN (Fig. 1g, Extended Data Fig. 1i), verifying the robust pathogenic effects of the *Ptpn11*<sup>E76K/+</sup> mutation in *Nestin*<sup>+</sup> BM stromal cells. These results suggested that the *Ptpn11*-mutated BM microenvironment drove MPN development by hyperactivation of resident wild-type HSCs. This notion was further supported by the observation that aberrant HSC activation occurred before full development of MPN in *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice (Extended Data Fig. 1j).

As *PTPN11* mutations in Noonan syndrome are present ubiquitously, we next determined the effect of the *Ptpn11*-mutated microenvironment on HSCs that also carried *Ptpn11* mutations. We compared *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> mice, in which Cre was expressed in haematopoietic cells as well as BM stromal cells<sup>10,11</sup> following administration of polyinosinic-polycytidylic acid (pI-pC), with *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> mice, in which constitutive Cre expression

The abundance of the neo cassette in genomic DNA was determined by qPCR ( $n = 3$  mice per group). **e–h**, BM cells collected from wild-type Boyl mice were transplanted into *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup>, *Ptpn11*<sup>+/+</sup>*Mx1-Cre*<sup>+</sup> (8 weeks following pI-pC treatment), and *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> (16 weeks old) mice. Recipients were monitored for MPN development for 6–8 months (**e**). Mac-1<sup>+</sup>Gr-1<sup>+</sup> myeloid cells in the BM, spleen and liver were examined ( $n = 5$  mice per group) (**f**). The pool size ( $n = 4$  mice per group) (**g**) and intracellular signalling activities ( $n = 3$  mice per group) (**h**) of donor HSCs were determined 25 weeks following transplantation. Data shown in **a**, **b**, **d**, **f–h** are mean  $\pm$  s.d. of all mice examined. Statistical significance was determined between *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> and *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> groups; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . Source Data for this figure are available online.

was restricted to haematopoietic cells and part of endothelial cells (see below). The disease phenotypes of *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> mice were much less severe (Fig. 2a, b, Extended Data Fig. 2a). Furthermore, at a 12-month follow-up check, MPN in 19% of *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> mice developed into acute leukaemia, as opposed to 63% of *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> mice (Fig. 2c). The inhibitory neo cassette in the mutated *Ptpn11* allele was deleted from haematopoietic cells to the same extent in both lines of mice. However, neo deletion from MSPCs, osteoblasts and endothelial cells was detected in *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> but not *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> mice (except for partial deletion from endothelial cells) (Fig. 2d). The differences in the severity and prognosis of MPN between these two lines of mice do not appear to be associated with pI-pC administration or the times/stages when the disease mutations were induced (Extended Data Fig. 2b–e). Furthermore, no donor-cell-derived MPN developed in *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> mice transplanted with wild-type BM cells, in contrast to the 75% incidence of donor-cell-derived MPN in *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> recipients (Fig. 2e, f, Extended Data Fig. 3a, b). Wild-type donor HSCs were also highly activated in *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup>, but not *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> recipients owing to aberrantly enhanced cell signalling activities (Fig. 2g, h). Similar results were obtained from the Noonan syndrome mutation *Ptpn11*<sup>D61G</sup> global knock-in mice, which were born with a developmental disorder resembling Noonan syndrome and developed JMML-like MPN<sup>4</sup>. Transplantation of wild-type BM cells into lethally-irradiated *Ptpn11*<sup>D61G/+</sup> mice initially reversed MPN. The mice appeared to be cured during the first 3 months after transplantation, but 8 out of 14 then developed donor-cell-derived MPN in the next 5 months (Extended Data Fig. 3c).

To further define the cell types in the *Ptpn11*-mutated BM microenvironment that have an important role in driving/enhancing MPN development, we generated cell-type-specific *Ptpn11*<sup>E76K/+</sup> knock-in mice and monitored them for one and a half years. The *Ptpn11*<sup>E76K/+</sup>



**Table 1 | *Ptpn11*<sup>E76K/+</sup> mutation in MSPCs and osteoprogenitors, but not differentiated osteoblasts or endothelial cells, in the BM microenvironment induces MPN**

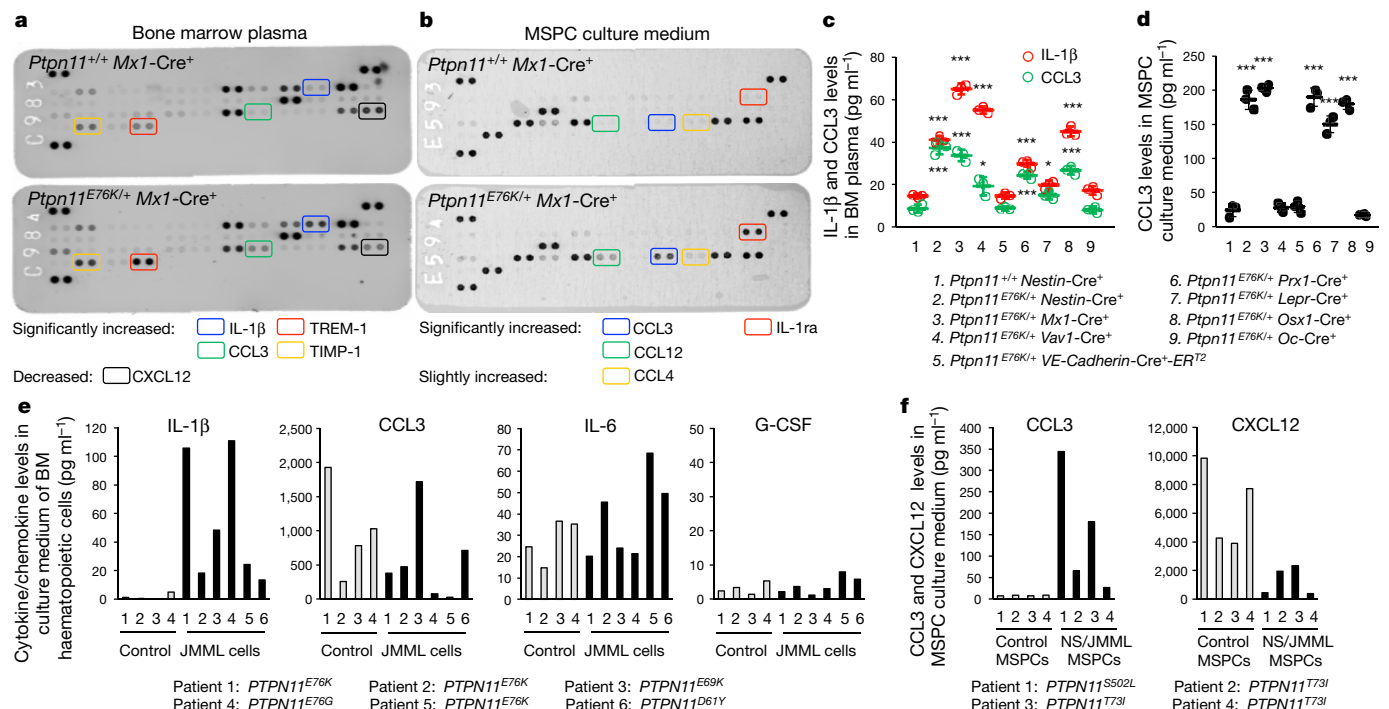
Cell-type-specific knock-in mice	Target cells	Age of mice euthanized	Incidence of MPN	HSC hyperactivation	Spleen weight (g)
<i>Ptpn11</i> <sup>E76K/+</sup> <i>Nestin-Cre</i> <sup>+</sup>	MSPCs	7–14 months	20/27***	Yes	0.289 ± 0.054
<i>Ptpn11</i> <sup>E76K/+</sup> <i>VE-Cadherin-Cre</i> <sup>+</sup> <i>-ERT2</i>	Endothelial cells	11–18 months	0/15	No	0.098 ± 0.056
<i>Ptpn11</i> <sup>E76K/+</sup> <i>Prx1-Cre</i> <sup>+</sup>	Mesenchymal cells	5–10 months	12/16***	Yes	0.385 ± 0.177
<i>Ptpn11</i> <sup>E76K/+</sup> <i>Lepr-Cre</i> <sup>+</sup>	Leptin receptor <sup>+</sup> mesenchymal cells	13–17 months	9/15***	Yes	0.281 ± 0.075
<i>Ptpn11</i> <sup>E76K/+</sup> <i>Osx1-Cre</i> <sup>+</sup>	Osteoprogenitors	5–8 months	13/14***	Yes	0.616 ± 0.08
<i>Ptpn11</i> <sup>E76K/+</sup> <i>Oc-Cre</i> <sup>+</sup>	Osteoblasts	11–18 months	0/16	No	0.109 ± 0.034

Cell-type-specific *Ptpn11*<sup>E76K</sup> knock-in mice as indicated were generated and monitored for MPN development for up to 18 months. The incidence of MPN, cycling status of BM HSCs, and spleen weights of the animals euthanized at the indicated ages were determined.

mutation in *Prx1*-expressing broad mesenchymal cells, *Lepr*<sup>+</sup> mesenchymal cells, Osterix (*Osx1*)-expressing osteoprogenitors (all of which contain/overlap with *Nestin*<sup>+</sup> MSPCs<sup>12–15</sup>), but not Osteocalcin (*Oc*)-expressing differentiated osteoblasts or VE-cadherin-expressing endothelial cells, induced MPN (Table 1, Extended Data Fig. 4a, b). The deletion efficiency of neo from mutated *Ptpn11* alleles in MSPCs generally correlated with the latency and severity of MPN that developed in these lines of cell-type-specific mutant mice (Extended Data Fig. 4c), suggesting that MSPCs and/or osteoprogenitors were responsible for the leukaemogenic effects of the *Ptpn11*-mutated BM microenvironment. HSCs were hyperactivated only in the lines of mice that developed MPN (Table 1), further underscoring the effect of HSC hyperactivation on the myeloid malignancy induced/enhanced by the *Ptpn11* mutation in MSPCs and osteoprogenitors.

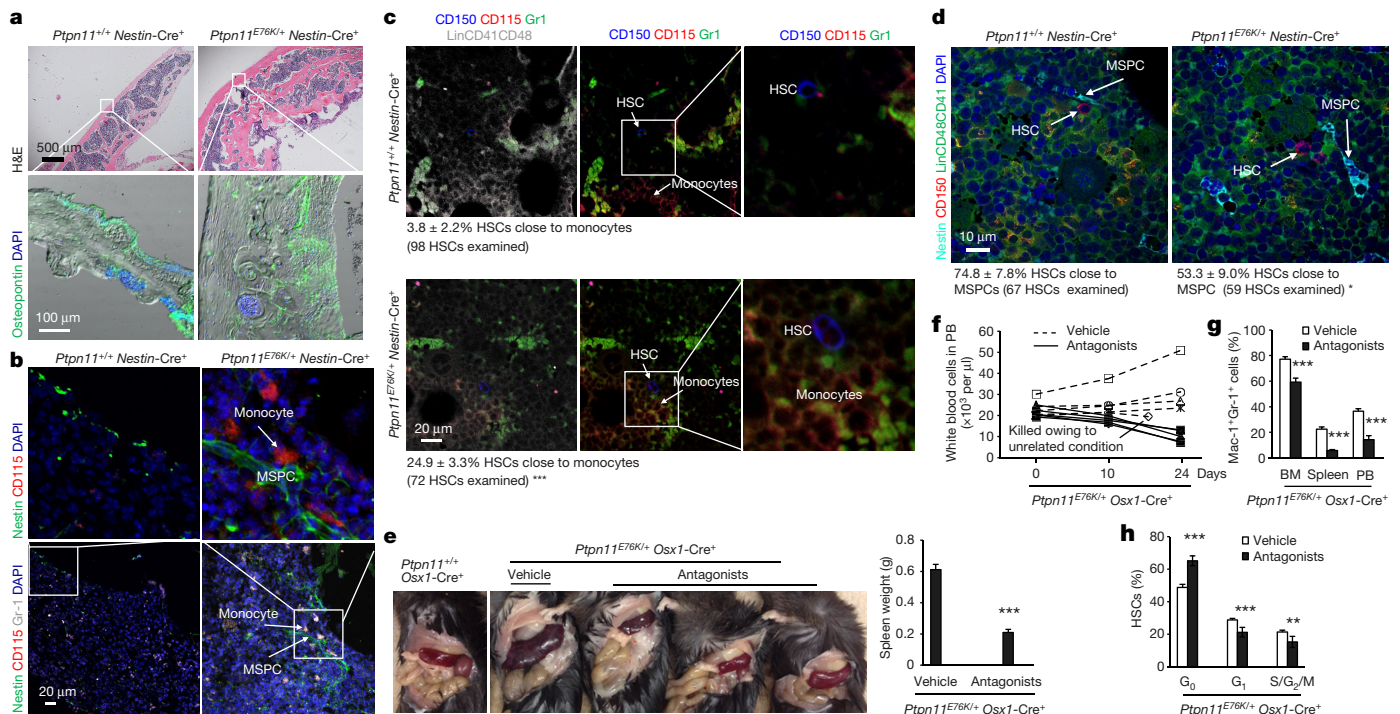
We next sought to identify the mechanisms by which *Ptpn11*-mutated MSPCs and osteoprogenitors activate HSCs (wild type or

mutant with the same *Ptpn11* mutation). Compared to wild-type HSCs, *Ptpn11*<sup>E76K/+</sup> mutant HSCs had accelerated myeloid differentiation owing to cell autonomous effects<sup>5</sup>, regardless of whether they were co-cultured with wild-type or *Ptpn11*<sup>E76K/+</sup> BM stromal cells or MSPCs (Extended Data Fig. 5a). Unexpectedly, *Ptpn11*<sup>E76K/+</sup> stromal cells and MSPCs had no significant activating effects on either *Ptpn11*<sup>E76K/+</sup> or wild-type HSCs (Extended Data Fig. 5a). Similar results were obtained when HSCs and MSPCs were co-cultured in two separate chambers that still allowed growth factors/cytokines to freely cross (Extended Data Fig. 5b). Interestingly, cytokine–chemokine array analyses for the BM plasma revealed that proinflammatory cytokines IL-1 $\beta$  and TREM-1, but not IL-6 (refs 16, 17), G-CSF<sup>18,19</sup>, GM-CSF<sup>16</sup>, TNF- $\alpha$ <sup>17,19,20</sup>, or IL-1 $\alpha$ <sup>17</sup> that are known to be involved in MPN, were substantially increased in *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> mice (Fig. 3a). In addition, the inflammatory CC chemokine CCL3 and TIMP-1, an inhibitor of matrix metalloproteinases generated by monocytes<sup>21</sup>, were increased,



**Figure 3 | IL-1 $\beta$  and CCL3 are overproduced by *Ptpn11*<sup>E76K/+</sup> MPN cells and MSPCs, respectively.** **a**, **b**, Chemokine–cytokine array analyses were performed with BM plasma collected from *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>*Mx1-Cre*<sup>+</sup> mice ( $n = 3$  mice per group) 12 weeks after pI-pC administration (**a**) and the culture medium of MSPCs derived from these mice (**b**). Representative results are shown. **c**, BM plasma collected from the indicated cell-type-specific *Ptpn11*<sup>E76K/+</sup> knock-in mice (7–8 months old;  $n = 3$  mice per group) was assayed for the levels of IL-1 $\beta$  and CCL3 by ELISA. **d**, CCL3 levels in the culture medium of MSPCs derived from the indicated lines of *Ptpn11*<sup>E76K/+</sup> knock-in mice ( $n = 3$  mice per group) were determined by ELISA. **e**, Levels of IL-1 $\beta$ , CCL3, IL-6 and G-CSF produced

by cells from patients with JMML with the indicated *PTPN11* mutations, and healthy human apheresis or BM cells were determined by cytometric bead array assay. **f**, CCL3 and CXCL12 levels in the culture medium of MSPCs derived from patients with Noonan syndrome (NS) with the indicated *PTPN11* mutations or healthy individuals were determined by cytometric bead array and ELISA assays, respectively. Each bar in **e**, **f** represents one patient or healthy individual. Data shown in **c**, **d** are mean  $\pm$  s.d. of all mice examined. Statistical significance was determined between the indicated cell-type-specific *Ptpn11*<sup>E76K/+</sup> knock-in mice and *Ptpn11*<sup>+/+</sup>*Nestin-Cre*<sup>+</sup> control mice; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . Source Data for this figure are available online.



**Figure 4** | *Ptpn11*<sup>E76K/+</sup> MSPCs recruit monocytes to the vicinity by excessive production of CCL3, leading to the hyperactivation and displacement of neighbouring HSCs from the niche. 7–10-month-old *Ptpn11*<sup>E76K/+</sup>Nestin-Cre<sup>+</sup> mice and *Ptpn11*<sup>+/+</sup>Nestin-Cre<sup>+</sup> littermates were analysed. **a**, Calvariums were processed for histopathological examination (haematoxylin and eosin (H&E) staining) and immunofluorescence staining with the indicated antibody. Slides were counterstained with the DNA dye DAPI (4',6-diamidino-2-phenylindole;  $n = 3$  mice per group). **b–d**, Bone sections (one section per femur or tibia) were immunostained with the indicated antibodies ( $n = 5$  mice per group). The spatial relationship between monocytes (CD115<sup>+</sup>Gr-1<sup>+</sup>) and MSPCs (Nestin<sup>+</sup>) (**b**), or between HSCs (Lin<sup>+</sup>CD48<sup>+</sup>CD41<sup>+</sup>CD150<sup>+</sup>) and monocytes (**c**) or MSPCs (**d**), were examined. HSCs within  $<8 \mu$ m of monocytes or MSPCs were considered as close to these cells. **e–h**, Six–seven-month-

whereas CXCL12 (SDF-1), a chemokine important for HSC retention in the niche<sup>22,23</sup>, was decreased (Fig. 3a). The spleen plasma from *Ptpn11*<sup>E76K/+</sup>*Mx1*-Cre<sup>+</sup> mice also showed markedly increased levels of CCL3, CCL12 and CCL4 (Extended Data Fig. 6a).

To comprehensively identify the protein factors that were aberrantly produced by MSPCs with *Ptpn11* mutations, we performed RNA-seq gene expression profiling analyses (GEO number GSE81311). mRNA levels of *Ccl3*, *Ccl12*, and *Ccl4* were increased by 6.5-, 3.7-, and 1.7-fold (log<sub>2</sub> scale), respectively, whereas expression of *Cxcl12* was decreased by 1.8-fold in *Ptpn11*<sup>E76K/+</sup> MSPCs (Extended Data Fig. 6b). In addition, the anti-inflammatory cytokine IL-1 receptor antagonist (*Il1ra*), was also increased by 4.6-fold. Cytokine–chemokine array analyses with MSPC culture medium confirmed that the amount of CCL3, CCL12, and IL-1ra proteins secreted by *Ptpn11*<sup>E76K/+</sup> MSPCs was indeed greatly increased (Fig. 3b). Levels of IL-1 $\beta$  and CCL3 in the BM plasma (Fig. 3c) and CCL3 in the culture medium of MSPCs (Fig. 3d) isolated from microenvironmental cell-type-specific *Ptpn11*<sup>E76K/+</sup> mice correlated closely with the latency and incidence of MPN in these lines of mice. Remarkably, IL-1 $\beta$  production from *PTPN11*-mutated leukaemic cells from patients with JMML also increased by 7.9-fold to 65.7-fold over that of healthy donor cells (Fig. 3e). The amount of CCL3 produced by MSPCs derived from *PTPN11*-mutation-positive Noonan syndrome patients with JMML complications increased by 3.3-fold to 43.0-fold, whereas CXCL12 was decreased compared to those secreted by normal human MSPCs (Fig. 3f). The direct effects of these aberrantly produced cytokines/chemokines on HSCs were then

old *Ptpn11*<sup>E76K/+</sup>*Osx1*-Cre<sup>+</sup> mice were treated with the CCR1 antagonist BX471 and the CCR5 antagonist Maraviroc or vehicle control for 23 days. Mice were killed, and spleens were weighed ( $n = 5$  and 10 mice for the antagonist and vehicle groups, respectively) (**e**). White blood cell counts in the peripheral blood were determined at the indicated time points ( $n = 5$  mice per group, each line represents one mouse) (**f**). Myeloid cells (Mac-1<sup>+</sup>Gr-1<sup>+</sup>) in the BM, spleen, and peripheral blood (PB) ( $n = 5$  and 10 mice for the antagonist and vehicle groups, respectively) (**g**), and the cell cycle distribution of HSCs in the BM ( $n = 5$  and 6 mice for the antagonist and vehicle groups, respectively) (**h**) were determined at the end of the experiments. Representative images are shown in **a–d**. Data shown in **e**, **g**, and **h** are mean  $\pm$  s.d. of all mice examined; \* $P < 0.01$ ; \*\*\* $P < 0.001$ . Source Data for this figure are available online.

determined. Interestingly, although IL-1 $\beta$  robustly activated HSCs to differentiate towards myeloid cells and monocytes, CCL3, CCL4, and CCL12—which were over-produced by *Ptpn11*-mutated MSPCs—did not show any activating effects on HSCs (Extended Data Fig. 6c).

We next investigated the *in vivo* consequences of the excessive CC chemokines produced by *Ptpn11*-mutated MSPCs. Nestin<sup>+</sup> MSPCs and Osteopontin<sup>+</sup> osteoblasts were increased in *Ptpn11*<sup>E76K/+</sup>Nestin-Cre<sup>+</sup> mice (Extended Data Fig. 7a). Frequencies of colony-forming unit fibroblasts (CFU-F) in the BM, indicative of MSPCs, were increased to various extents (Extended Data Fig. 7b) that were commensurate with the induction efficiencies of the *Ptpn11*<sup>E76K/+</sup> mutation in MSPCs in various lines of microenvironmental cell-type-specific knock-in mice (Extended Data Fig. 4c). Indeed, MSPCs isolated from *Ptpn11*<sup>E76K/+</sup>Nestin-Cre<sup>+</sup> mice grew much faster with significantly enhanced cycling due to elevated cell signalling activities caused by the activating mutation of Shp2 (Extended Data Fig. 7c–e). In addition, osteogenesis was enhanced in *Ptpn11*<sup>E76K/+</sup>Nestin-Cre<sup>+</sup> mice as evidenced by markedly increased thickness of the calvarium (Fig. 4a). Most notably, Nestin<sup>+</sup> MSPCs in *Ptpn11*<sup>E76K/+</sup>Nestin-Cre<sup>+</sup> (Fig. 4b) and *Ptpn11*<sup>E76K/+</sup>*Prx1*-Cre<sup>+</sup> mice (Extended Data Fig. 8a) were frequently surrounded by CD115<sup>+</sup>Gr-1<sup>+</sup> inflammatory monocytes, but not F4/80<sup>+</sup> macrophages (Extended Data Fig. 8b). This was probably attributable to the excessive CCL3 and possibly other CC chemokines secreted from *Ptpn11*<sup>E76K/+</sup> MSPCs, because these chemokines strongly induce chemotaxis of monocytes<sup>24,25</sup>. Consequently, the percentage of HSCs surrounded by CD115<sup>+</sup>Gr-1<sup>+</sup> monocytes greatly increased (Fig. 4c)



and the percentage of HSCs close to Nestin<sup>+</sup> MSPCs decreased (Fig. 4d). Furthermore, the distance of HSCs from CD31<sup>+</sup>CD144<sup>+</sup> endothelial cells doubled (Extended Data Fig. 8c) and the percentage of HSCs residing in the megakaryocyte niches significantly decreased (Extended Data Fig. 8d) in *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> mice. Thus, it appears that persistent high levels of proinflammatory cytokines produced by the monocytes (with or without the *Ptpn11* mutation) recruited by *Ptpn11*-mutated MSPCs/osteoprogenitors hyperactivated neighbouring HSCs with the same mutation or wild-type donor HSCs and displaced them from MSPC, endothelial cell, and megakaryocyte niches that are essential for maintaining HSC dormancy<sup>8,9,26–29</sup>, resulting in exacerbated MPN or donor-cell-derived MPN.

To validate the role of excessive CCL3 in mediating the pathogenic effects of the *Ptpn11*-mutated BM microenvironment, we treated *Ptpn11*<sup>E76K/+</sup>*Osx1-Cre*<sup>+</sup> mice with the CCL3 receptor (CCR1 and CCR5) antagonists. As shown in Figure 4e–g and Extended Data Figure 9a, b, treatment with CCR1 and CCR5 antagonists for 3 weeks effectively reversed MPN phenotypes, as determined by spleen weights, total white blood cell counts in the peripheral blood, and myeloid cells in the BM, spleen, and peripheral blood. The therapeutic effects correlated with the restoration of the quiescence and the size of the HSC pool (Fig. 4h, Extended Data Fig. 9c). We also treated *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> mice with CCL3 receptor antagonists. Similar effects, but to a lesser extent, were observed (Extended Data Fig. 10a–e).

In summary, our mouse genetics studies have demonstrated that *Ptpn11* mutations in the BM microenvironment have pathogenic effects on resident HSCs, promoting/inducing leukaemogenesis. Nevertheless, as Noonan syndrome involves various mutations in *PTPN11* and other genes (such as *RAS*, *CBL*, *B-RAF*, *SOS1*, and *SHOC2*), it remains to be determined whether the leukaemogenic effects of microenvironmental *PTPN11* mutations depend on the potencies of these mutations, and whether Noonan-syndrome-associated mutations in other genes in the BM microenvironment also have detrimental effects. Clinical phenotype-genotype correlative studies in a large cohort of Noonan syndrome patients are required to address these questions.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 27 October 2015; accepted 29 September 2016.**

**Published online 26 October 2016.**

- Mohi, M. G. & Neel, B. G. The role of Shp2 (*PTPN11*) in cancer. *Curr. Opin. Genet. Dev.* **17**, 23–30 (2007).
- Tartaglia, M. *et al.* Mutations in *PTPN11*, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat. Genet.* **29**, 465–468 (2001).
- Roberts, A. E., Allanson, J. E., Tartaglia, M. & Gelb, B. D. Noonan syndrome. *Lancet* **381**, 333–342 (2013).
- Araki, T. *et al.* Mouse model of Noonan syndrome reveals cell type- and gene dosage-dependent effects of *Ptpn11* mutation. *Nat. Med.* **10**, 849–857 (2004).
- Xu, D. *et al.* Non-lineage/stage-restricted effects of a gain-of-function mutation in tyrosine phosphatase *Ptpn11* (Shp2) on malignant transformation of hematopoietic cells. *J. Exp. Med.* **208**, 1977–1988 (2011).
- Chan, G. *et al.* Leukemogenic *Ptpn11* causes fatal myeloproliferative disorder via cell-autonomous effects on multiple stages of hematopoiesis. *Blood* **113**, 4414–4424 (2009).
- Mohi, M. G. *et al.* Prognostic, therapeutic, and mechanistic implications of a mouse model of leukemia evoked by Shp2 (*PTPN11*) mutations. *Cancer Cell* **7**, 179–191 (2005).
- Méndez-Ferrer, S. *et al.* Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature* **466**, 829–834 (2010).
- Kunisaki, Y. *et al.* Arterial niches maintain haematopoietic stem cell quiescence. *Nature* **502**, 637–643 (2013).
- Park, D. *et al.* Endogenous bone marrow MSCs are dynamic, fate-restricted participants in bone maintenance and regeneration. *Cell Stem Cell* **10**, 259–272 (2012).

- Walkley, C. R., Shea, J. M., Sims, N. A., Purton, L. E. & Orkin, S. H. Rb regulates interactions between hematopoietic stem cells and their bone marrow microenvironment. *Cell* **129**, 1081–1095 (2007).
- Kfoury, Y. & Scadden, D. T. Mesenchymal cell contributions to the stem cell niche. *Cell Stem Cell* **16**, 239–253 (2015).
- Mendelson, A. & Frenette, P. S. Hematopoietic stem cell niche maintenance during homeostasis and regeneration. *Nat. Med.* **20**, 833–846 (2014).
- Morrison, S. J. & Scadden, D. T. The bone marrow niche for haematopoietic stem cells. *Nature* **505**, 327–334 (2014).
- Mizoguchi, T. *et al.* Osterix marks distinct waves of primitive and definitive stromal progenitors during bone marrow development. *Dev. Cell* **29**, 340–349 (2014).
- Reynaud, D. *et al.* IL-6 controls leukemic multipotent progenitor cell fate and contributes to chronic myelogenous leukemia development. *Cancer Cell* **20**, 661–673 (2011).
- Zhang, B. *et al.* Altered microenvironmental regulation of leukemic and normal stem cells in chronic myelogenous leukemia. *Cancer Cell* **21**, 577–592 (2012).
- Schepers, K. *et al.* Myeloproliferative neoplasia remodels the endosteal bone marrow niche into a self-reinforcing leukemic niche. *Cell Stem Cell* **13**, 285–299 (2013).
- Wang, L. *et al.* Notch-dependent repression of miR-155 in the bone marrow niche regulates hematopoiesis in an NF-κB-dependent manner. *Cell Stem Cell* **15**, 51–65 (2014).
- Walkley, C. R. *et al.* A microenvironment-induced myeloproliferative syndrome caused by retinoic acid receptor gamma deficiency. *Cell* **129**, 1097–1110 (2007).
- Khokha, R., Murthy, A. & Weiss, A. Metalloproteinases and their natural inhibitors in inflammation and immunity. *Nat. Rev. Immunol.* **13**, 649–665 (2013).
- Ding, L. & Morrison, S. J. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* **495**, 231–235 (2013).
- Greenbaum, A. *et al.* CXCL12 in early mesenchymal progenitors is required for haematopoietic stem-cell maintenance. *Nature* **495**, 227–230 (2013).
- Shi, C. & Pamer, E. G. Monocyte recruitment during infection and inflammation. *Nat. Rev. Immunol.* **11**, 762–774 (2011).
- Youn, B. S., Mantel, C. & Broxmeyer, H. E. Chemokines, chemokine receptors and hematopoiesis. *Immunol. Rev.* **177**, 150–174 (2000).
- Bruns, I. *et al.* Megakaryocytes regulate hematopoietic stem cell quiescence through CXCL4 secretion. *Nat. Med.* **20**, 1315–1320 (2014).
- Zhao, M. *et al.* Megakaryocytes maintain homeostatic quiescence and promote post-injury regeneration of hematopoietic stem cells. *Nat. Med.* **20**, 1321–1326 (2014).
- Kiel, M. J. *et al.* SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**, 1109–1121 (2005).
- Ding, L., Saunders, T. L., Enikolopov, G. & Morrison, S. J. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature* **481**, 457–462 (2012).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was supported by The National Institutes of Health grants HL130995 and DK092722 (to C.K.Q.).

**Author Contributions** L.D. generated microenvironmental cell-type-specific knock-in mice, analysed MPN development and progression, performed HSC imaging and MSPC signalling analyses, and tested the therapeutic effects of CCL3 receptor antagonists. W.-M.Y. trained L.D. and H.Z. in techniques, performed cytokine array analyses, and analysed patient specimens. H.Z. generated MSPC-specific knock-in mice, characterized MPN development and progression, and analysed HSC phenotypes. S.T.B. and M.P. identified and collected patient specimens. M.L.L. provided patient specimens and thoroughly discussed the work. M.Z., G.H., H.E.B., and D.T.S. provided critical advice on experimental designs and interpretation of the data, and edited the manuscript. C.-K.Q. designed the experiments and directed the entire project. L.D. and C.K.Q. wrote the manuscript with input from all authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.-K. Qu ([cheng-kui.qu@emory.edu](mailto:cheng-kui.qu@emory.edu)).

**Reviewer Information** Nature thanks I. Ghobrial, B. Neel and the other anonymous reviewer(s) for their contribution to the peer review of this work.



## METHODS

**Mice.** Generation of *Ptpn11*<sup>E76K-neo/+</sup> mice have previously been reported<sup>5</sup>. A neo cassette with a stop codon flanked by *loxP* sites was inserted in the second intron of the *Ptpn11* allele followed by the mutation GAA (E) to AAA (K) at the amino acid 76 encoding position in the third exon. The mice were backcrossed to C57BL/6 mice for more than 10 generations. *Ptpn11*<sup>D61G/+</sup> mice<sup>4</sup> were originally imported from Beth Israel Deaconess Medical Center. *Nestin-Cre*<sup>+30</sup>, *Mx1-Cre*<sup>+31</sup>, *Vav1-Cre*<sup>+32</sup>, *Prx1-Cre*<sup>+33</sup>, *Lepr-Cre*<sup>+34</sup>, *Osx1-Cre*<sup>+35</sup>, *Oc-Cre*<sup>+36</sup>, and *VE-Cadherin-Cre*<sup>-ER<sup>T2</sup></sup> (ref. 37) transgenic mice used in this study were purchased from the Jackson Laboratory or obtained from the investigators who originally developed the mouse lines. Mice of the same age, sex, and genotype were mixed and then randomly grouped for subsequent analyses (investigators were not blinded during allocation, during experiments and outcome assessment). All mice were kept under specific-pathogen-free conditions in the Animal Resources Center at Case Western Reserve University and subsequently Emory University Division of Animal Resources. All animal procedures complied with the NIH Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committee.

*Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> mice and *Ptpn11*<sup>+/+</sup>*Mx1-Cre*<sup>+</sup> littermates (8 weeks old) were administered with i.p. injection of 3 doses of pI-pC (1.0 µg per g body weight) every other day over 5 days. *Ptpn11*<sup>E76K/+</sup>*VE-Cadherin-Cre*<sup>+</sup>*ER<sup>T2</sup>* mice and *Ptpn11*<sup>+/+</sup>*VE-Cadherin-Cre*<sup>+</sup>*ER<sup>T2</sup>* littermates (4–6 weeks old) were administered with i.p. injection of 3 doses of tamoxifen (9.0 mg per 40 g body weight) every other day over 5 days. Mice were analysed at the indicated time points after pI-pC or tamoxifen administration. Acute leukaemia progression in pI-pC administered *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> and *Ptpn11*<sup>E76K/+</sup>*Vav1-Cre*<sup>+</sup> mice was determined as we previously described<sup>5</sup>. No statistical methods were used to predetermine sample size.

**Patient specimens.** De-identified BM biopsies from *PTPN11*-mutation-positive Noonan syndrome patients with JMML or non-syndromic *PTPN11* mutation-positive patients with JMML were obtained from the University of California, San Francisco Tissue Cancer Cell Bank and Children's Healthcare of Atlanta, Emory University. Informed consent was obtained from all subjects. The experiments involving human subjects were reviewed and approved (Exemption IV) by the Institutional Review Board of Emory University.

**BM cell transplantation.** BM cells ( $2 \times 10^6$ ) collected from indicated donor mice were transplanted into lethally irradiated (1,100 cGy) recipient mice with the indicated genotypes through tail vein injection. Recipients were monitored for MPN development for 6–8 months.

**Quantitative real-time PCR (qPCR).** To determine the abundance of the neo cassette in the targeted *Ptpn11* allele, genomic DNA of haematopoietic cells, MSPCs, or other indicated cells was extracted with a ZR-Duet DNA/RNA MiniPrep extraction kit (Zymo Research). The abundance of the neo cassette was then quantified by qPCR using the Applied Biosystems 7500 Fast Real-Time PCR System. The PCR primers used were: 5'-TGGGAAGACAATAGCAGGCA-3' and 5'-CCCACTCA CCTTGTCATGTA-3'.

**Fluorescence-activated cell sorting (FACS).** The pool size, cell cycle status, apoptosis, and cell signalling activities of HSCs were analysed by multiparameter FACS analyses, as previously described<sup>38</sup>. In brief, for the HSC-pool-size analysis, fresh BM cells were stained with the following antibodies (eBiosciences, San Diego, unless otherwise noted): lineage antibodies (B220 (RA3-6B2), CD3 (145-2C11), Gr-1 (RB6-8C5), Mac-1 (M1/70), and Ter-119 (TER-119)), anti-Sca-1 (D7, BD Biosciences), anti-c-Kit (2B8), anti-CD150 (TC15-12F12.2, BD Biosciences), anti-CD48 (HM48-1), and anti-Flk2 (A2F10.1). Lin<sup>-</sup>Sca-1<sup>+</sup>c-Kit<sup>+</sup>CD150<sup>+</sup>CD48<sup>+</sup>Flk2<sup>-</sup> cells were quantified as HSCs. For the cell cycle analysis, freshly collected BM cells were stained for HSCs as above. Cells were then fixed and permeabilized using a Cytofix/Cytoperm kit (BD Biosciences), stained with Ki-67 antibody, and further incubated with Hoechst 33342 (20 µg ml<sup>-1</sup>). For the apoptosis analysis, BM cells were stained for HSCs, and then incubated with Annexin V and 7-amino-actinomycin D (BD Biosciences). For cell signalling analyses, BM cells were stained for HSCs, fixed and permeabilized using a Cytofix/Cytoperm kit, and then stained with anti-phospho-Erk (mouse IgG) (E-4, Santa Cruz Biotechnology), anti-phospho-Akt (rabbit IgG) (C31E5E, Cell Signaling), or anti-phospho-NF-κB (rabbit IgG) (93H1, Cell Signaling) antibodies, washed and further incubated with AlexaFluor488-conjugated secondary antibodies (goat anti-mouse IgG or goat anti-rabbit IgG) (Life technologies). Phosphorylation levels of these signalling proteins were determined by mean fluorescence intensities (MFI) of gated cells. Data were collected on BD LSR II Flow Cytometer (BD Biosciences) and analysed with FlowJo (Treestar).

**In vitro HSC culture.** HSCs (Lin<sup>-</sup>Sca-1<sup>+</sup>c-Kit<sup>+</sup>CD150<sup>+</sup>CD48<sup>+</sup>Flk2<sup>-</sup>) sorted from wild-type C57BL/6 mice were cultured in StemSpan medium supplemented with SCF (50 ng ml<sup>-1</sup>), Flt3 ligand (50 ng ml<sup>-1</sup>), TPO (50 ng ml<sup>-1</sup>), IL-3 (20 ng ml<sup>-1</sup>),

and IL-6 (20 ng ml<sup>-1</sup>) in the presence of IL-1β (10 ng ml<sup>-1</sup>), CCL3 (20 ng ml<sup>-1</sup>), CCL4 (20 ng ml<sup>-1</sup>), or CCL12 (20 ng ml<sup>-1</sup>). Six days later, cells were collected and analysed for Mac-1<sup>+</sup> myeloid cells, F4/80<sup>+</sup> macrophages, and CD115<sup>+</sup> monocytes.

**MSPC isolation and enrichment.** Mouse MSPCs were enriched following a standard protocol<sup>39</sup>. In brief, BM was collected from long bones. The bones were then crushed and digested with collagenase type II (2.5 mg ml<sup>-1</sup>) (Worthington Biochemical Corporation). BM cells and digested bone fragments were combined and cultured in DMEM supplemented with 15% fetal bovine serum (FBS). For human MSPC derivation, only BM cells were used. Suspension haematopoietic cells were removed after 24 h. Medium was replenished every 72 h. Colonies of MSPCs appeared 6–8 days after initial plating. To further purify MSPCs, cells were collected and stained with biotin-conjugated CD45 antibody and anti-biotin microbeads. CD45<sup>+</sup> haematopoietic cells were depleted using MACS separation columns (Miltenyi Biotec Inc.). The purity of MSPCs (>95%) was further confirmed according to the (CD45<sup>+</sup>CD140α<sup>+</sup>Sca-1<sup>+</sup>) phenotypes<sup>39</sup> by multiparameter FACS analyses.

**Fibroblast colony-forming unit (CFU-F) and colony forming unit-granulocyte/macrophage (CFU-GM) assays.** For the CFU-F assay,  $2 \times 10^6$  unfractionated BM cells were plated and cultured for 10–14 days as described above. Cells were stained with 0.5% crystal violet (Sigma-Aldrich) in 10% methanol for 20 min. Colonies formed by more than 50 fibroblast-like cells were counted under a light microscope. For the CFU-GM assay, freshly collected BM cells ( $2 \times 10^4$  cells ml<sup>-1</sup>) were seeded in 0.9% methylcellulose IMDM medium containing 30% FBS, glutamine (10<sup>-4</sup> M), β-mercaptoethanol (3.3 × 10<sup>-5</sup> M), and IL-3 (1 ng ml<sup>-1</sup>) or GM-CSF (1 ng ml<sup>-1</sup>). After 7 days of culture at 37 °C in a humidified 5% CO<sub>2</sub> incubator, colonies (primarily CFU-GM) formed by more than 50 haematopoietic cells were counted under an inverted microscope.

**RNA-sequencing analysis.** MSPCs (CD45<sup>+</sup>Ter-119<sup>-</sup>CD31<sup>-</sup>CD140α<sup>+</sup>Sca-1<sup>+</sup>)<sup>39</sup> were freshly isolated from the BM of *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>*Nestin-Cre*<sup>+</sup> mice. RNA was extracted using the RNeasy Midi kit (Qiagen). Total RNA samples were enriched for polyadenylated transcripts using the Oligotex mRNA Mini kit (Qiagen), and strand-specific RNA-seq libraries were generated using PrepX RNA library preparation kits (IntegenX), following the manufacturer's protocol. After cleanup with AMPure XP beads (Beckman Coulter) and amplification with Phusion High-Fidelity polymerase (New England BioLabs), RNA libraries were sequenced on a HiSeq 4000 instrument to a depth of at least 20 million reads. The correlation coefficient between the two groups is 0.954, which verifies that the method is accurate (Extended Data Fig. 6b). Before differential gene expression analysis, for each sequenced library, the read counts were adjusted by edgeR program package through one scaling normalized factor. Differential expression analysis of two conditions was performed using the DESeq R package (1.12.0). The *P* values were adjusted using the Benjamini–Hochberg method. Corrected *P* value of 0.005 and log<sub>2</sub>(fold change) of 1 were set as the threshold for significantly different expression.

**Chemokine–cytokine array analyses.** Femurs were dissected from *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> mice and *Ptpn11*<sup>+/+</sup>*Mx1-Cre*<sup>+</sup> littermates 12 weeks after pI-pC administration. BM plasma was collected by flushing one femur with 1.0 ml of phosphate buffered saline (PBS). MSPCs derived from pI-pC-administered *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>*Mx1-Cre*<sup>+</sup> mice were cultured ( $4 \times 10^6$  cells in 2.0 ml medium) in serum-free DMEM for 48 h. The culture medium was then collected. BM plasma or MSPC culture medium were analysed with Mouse Cytokine Antibody Array blots (R&D Systems) following the instructions provided by the manufacturer.

**ELISA and cytometric bead array assay.** BM plasma collected from one femur and one tibia in 500 µl PBS. Culture medium was collected from mouse MSPCs ( $4 \times 10^6$  cells per 2.0 ml) at second or third passages cultured in serum-free DMEM for 48 h. These samples were assayed for levels of IL-1β and CCL3 using enzyme-linked immunosorbent assay (ELISA) kits (IL-1β: eBioscience; CCL3: R&D Systems) following the instructions provided by the manufacturers. To determine multiple cytokines/chemokines produced by human MSPCs, MSPCs ( $2 \times 10^4$  cells ml<sup>-1</sup>) were cultured in serum-free StemSpan medium for 72–96 h. To determine multiple protein factors produced by cells from patients with JMML, JMML cells ( $2 \times 10^5$  cells ml<sup>-1</sup>) were cultured in StemSpan medium supplemented with human SCF (50 ng ml<sup>-1</sup>), human Flt3 ligand (50 ng ml<sup>-1</sup>), and human TPO (50 ng ml<sup>-1</sup>) for 72 h. The culture medium was then collected and cytokine/chemokine levels were determined by the BD Cytometric Bead Array Flex Sets (BD Biosciences) following the manufacturer's instructions. Human CXCL12 levels in MSPC culture medium were measured using a Human CXCL12/SDF-1 alpha Quantikine ELISA Kit (R&D systems).

**Immunofluorescence staining.** Frozen tissue sections prepared from 4% paraformaldehyde-fixed and decalcified bones were thawed at room temperature and then rehydrated with PBS. The slides were stained with the following antibodies

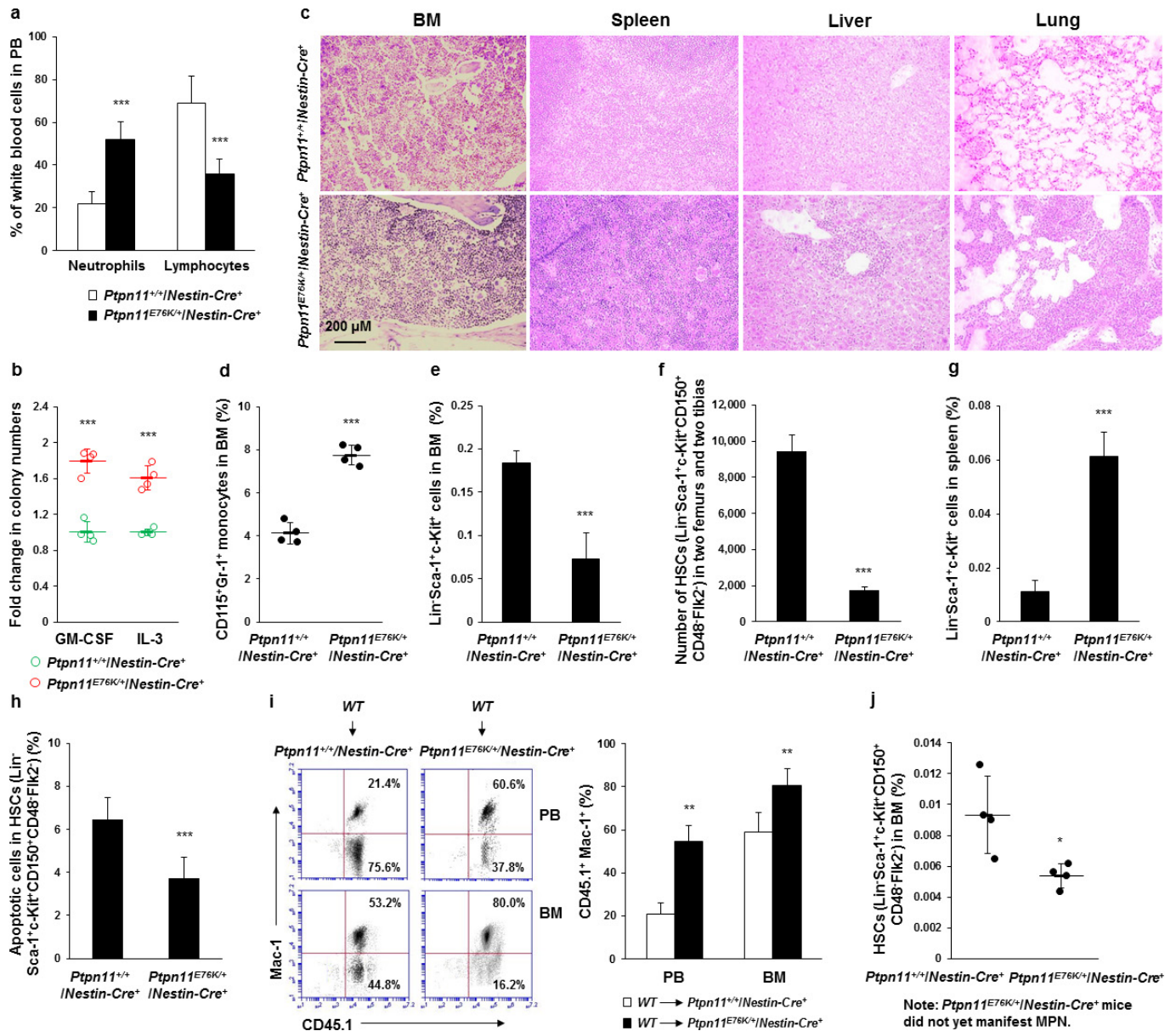
(eBiosciences, San Diego, unless otherwise noted) following standard procedures: anti-Osteopontin (Abcam), anti-Nestin (MAB353, Millipore), anti-Gr-1 (RB6-8C5), anti-Mac-1 (M1/70), anti-B220 (RA3-6B2), anti-Ter-119 (TER-119), anti-CD3 (145-2C11, BD Biosciences), anti-CD115 (AFS98), anti-CD150 (TC15-12F12.2, BD Biosciences), anti-CD31 (MEC13.3, Biolegend), anti-CD48 (HM48-1), and anti-CD41 (eBioMWRReg30) antibodies. Images were acquired using Olympus Confocal Laser Scanning Biological Microscope FV1000 equipped with four lasers ranging from 405 to 635 nm. Images were processed with ImageJ software.

**Administration of CCR1 and CCR5 antagonists.** *Ptpn11*<sup>E76K/+</sup>*Osx1*-Cre<sup>+</sup> mice (6–7 month old) and *Ptpn11*<sup>E76K/+</sup>*Mx1*-Cre<sup>+</sup> mice (4 weeks after pI-pC administration) were treated daily via subcutaneous injection with the CCR1 antagonist BX471 ((2*R*)-1-((2-((aminocarbonyl)amino)-4-chlorophenoxy)acetyl)-4-((4-fluorophenyl)methyl)-2-methylpiperazine) purchased from Tocris Bioscience (50 mg kg<sup>-1</sup> of body weight). These animals also received the CCR5 antagonist Maraviroc (4,4-difluoro-*N*-((*S*)-3-(3-(3-isopropyl-5-methyl-4*H*-1,2,4-triazol-4-yl)-8-azabicyclo(3.2.1)octan-8-yl)-1-phenylpropyl)cyclohexanecarboxamide) obtained from Selleck Chemicals (0.3 mg ml<sup>-1</sup> in the drinking water). Control *Ptpn11*<sup>E76K/+</sup>*Osx1*-Cre<sup>+</sup> mice and *Ptpn11*<sup>E76K/+</sup>*Mx1*-Cre<sup>+</sup> mice were given vehicle (70% ethanol and 0.5% DMSO for subcutaneous injections, and 1% DMSO in drinking water). Mice were treated for 23 days and then killed for subsequent analyses.

**Statistics.** Data are presented as mean ± s.d. of all mice analysed in multiple experiments (that is, biological replicates). Statistical significance was determined using unpaired two-tailed Student's *t* test. For HSC imaging analyses, two-tier tests were used to first combine technical replicates and then evaluate biological replicates.

To determine statistical significance in the incidences of MPN development and malignant progression, Fisher's exact tests were performed. \**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001; N.S., not significant in Extended Data Figs 2, 5.

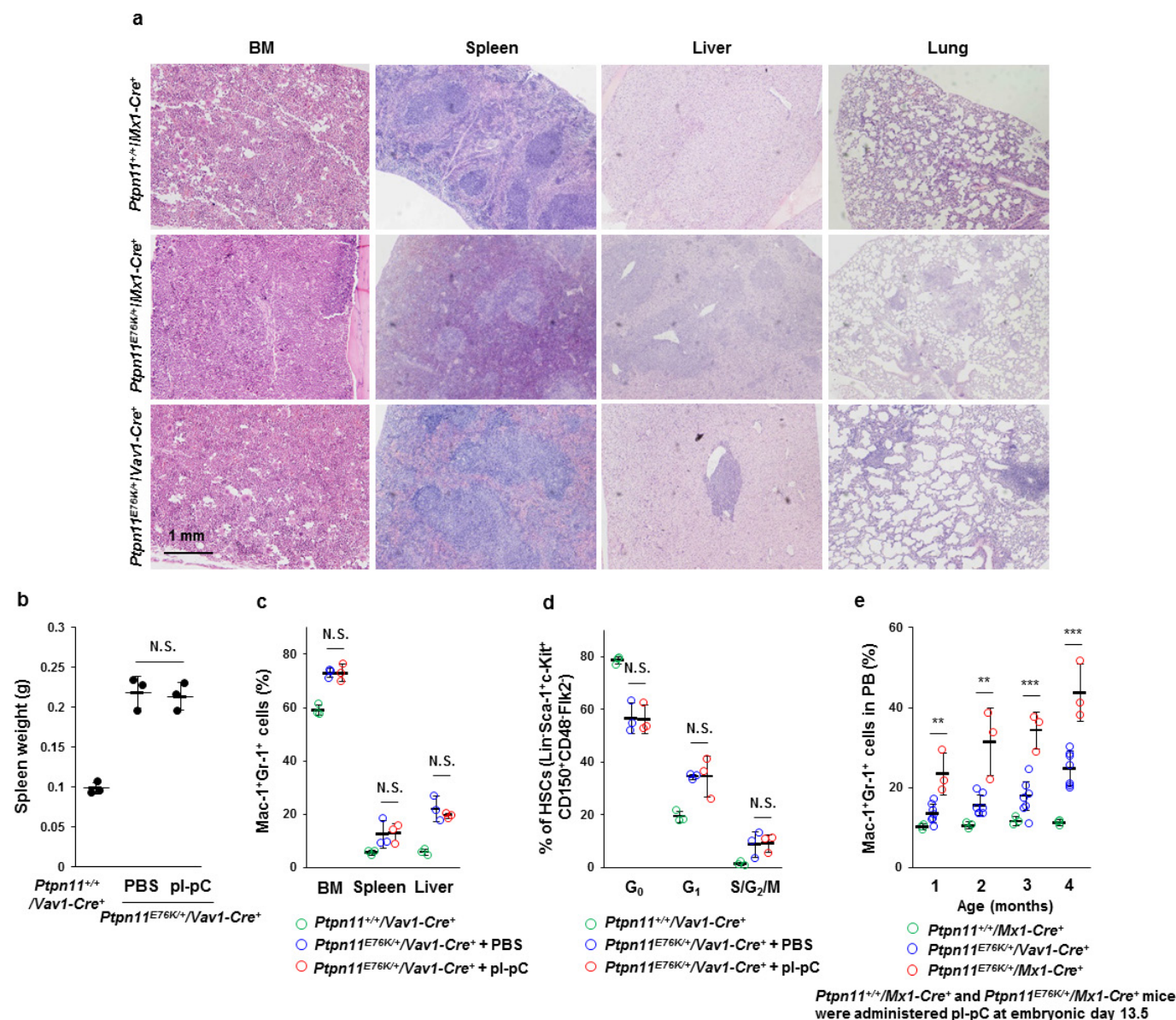
30. Tronche, F. *et al.* Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. *Nat. Genet.* **23**, 99–103 (1999).
31. Kühn, R., Schwenk, F., Aguet, M. & Rajewsky, K. Inducible gene targeting in mice. *Science* **269**, 1427–1429 (1995).
32. Stadtfeld, M. & Graf, T. Assessing the role of hematopoietic plasticity for endothelial and hepatocyte development by non-invasive lineage tracing. *Development* **132**, 203–213 (2005).
33. Logan, M. *et al.* Expression of Cre Recombinase in the developing mouse limb bud driven by a *Prxl* enhancer. *Genesis* **33**, 77–80 (2002).
34. DeFalco, J. *et al.* Virus-assisted mapping of neural inputs to a feeding center in the hypothalamus. *Science* **291**, 2608–2613 (2001).
35. Rodda, S. J. & McMahon, A. P. Distinct roles for Hedgehog and canonical Wnt signaling in specification, differentiation and maintenance of osteoblast progenitors. *Development* **133**, 3231–3244 (2006).
36. Zhang, M. *et al.* Osteoblast-specific knockout of the insulin-like growth factor (IGF) receptor gene reveals an essential role of IGF signaling in bone matrix mineralization. *J. Biol. Chem.* **277**, 44005–44012 (2002).
37. Monvoisin, A. *et al.* VE-cadherin-CreER<sup>T2</sup> transgenic mouse: a model for inducible recombination in the endothelium. *Dev. Dyn.* **235**, 3413–3422 (2006).
38. Yu, W. M. *et al.* Metabolic regulation by the mitochondrial phosphatase PTPMT1 is required for hematopoietic stem cell differentiation. *Cell Stem Cell* **12**, 62–74 (2013).
39. Morikawa, S. *et al.* Prospective identification, isolation, and systemic transplantation of multipotent mesenchymal stem cells in murine bone marrow. *J. Exp. Med.* **206**, 2483–2496 (2009).

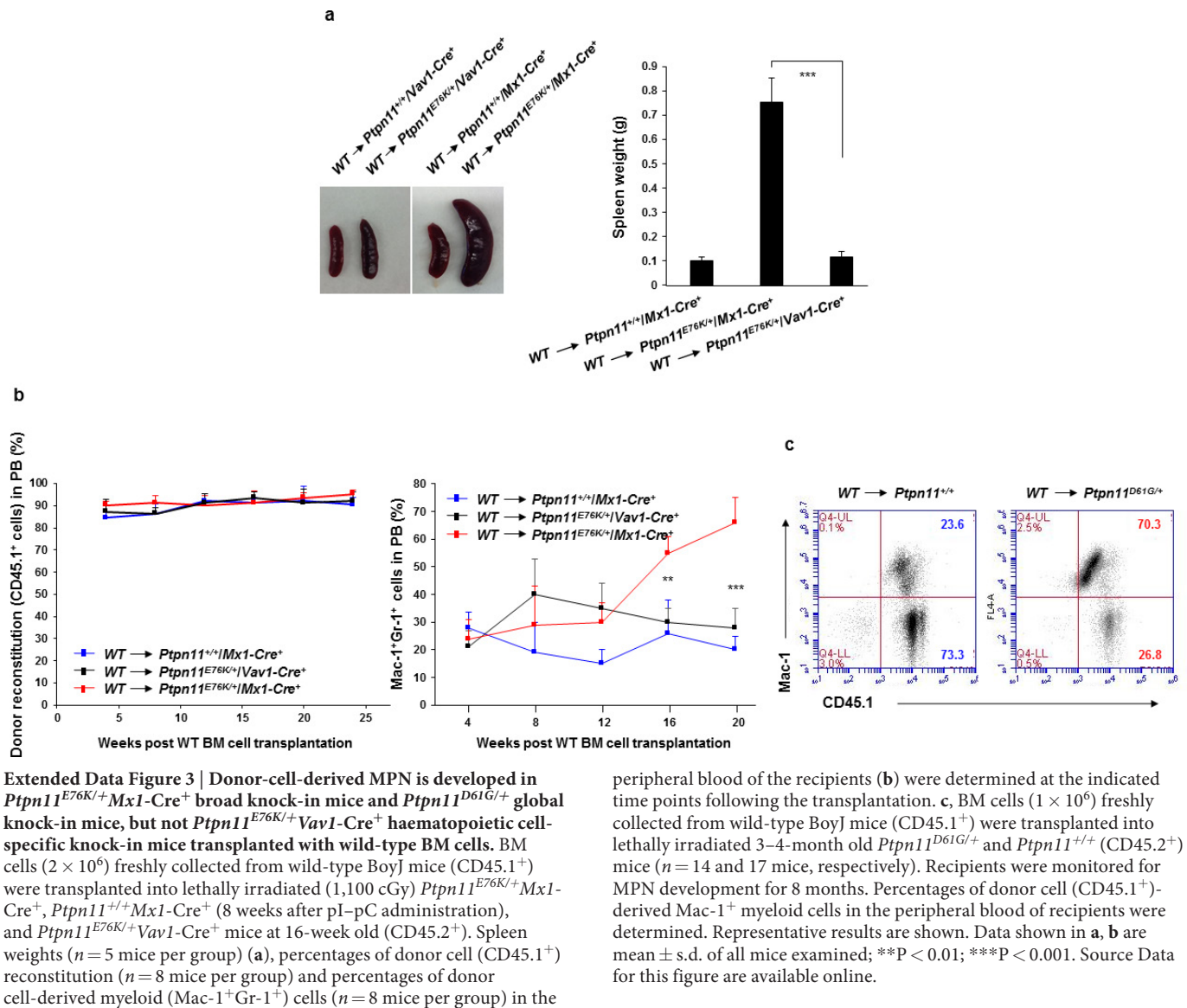


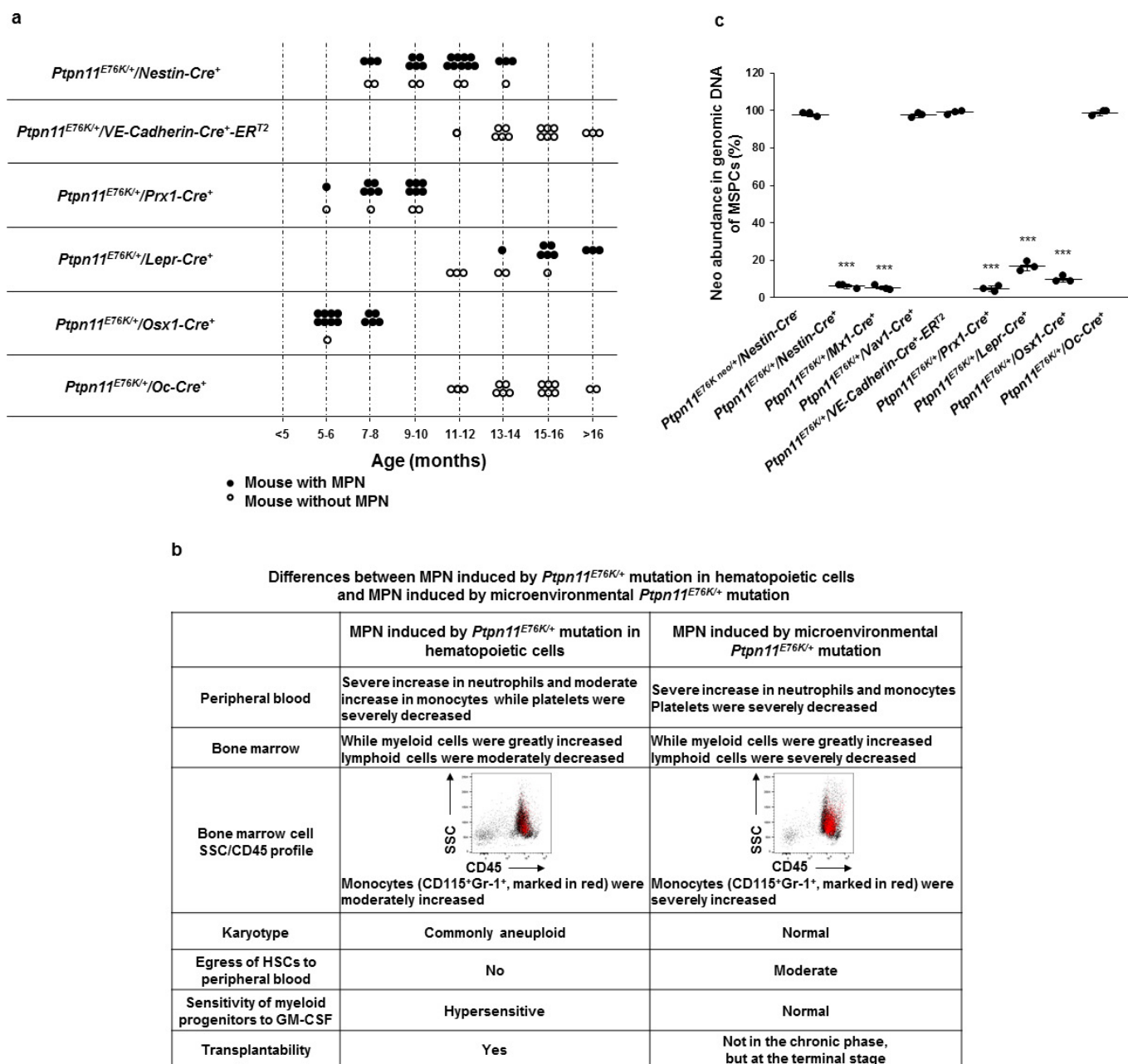
**Extended Data Figure 1 | *Ptpn11<sup>E76K/+</sup>* mutation in MSPCs induces MPN by aberrant activation of neighbouring wild-type HSCs in *Ptpn11<sup>E76K/+</sup>/Nestin-Cre<sup>+</sup>* mice.** **a**, Peripheral blood collected from 7–12-month-old *Ptpn11<sup>E76K/+</sup>/Nestin-Cre<sup>+</sup>* mice with MPN and *Ptpn11<sup>+/+</sup>/Nestin-Cre<sup>+</sup>* littermates were analysed for percentages of neutrophils and lymphocytes ( $n = 15$  mice per group). **b**, BM cells ( $2 \times 10^4$  cells) freshly collected from *Ptpn11<sup>E76K/+</sup>/Nestin-Cre<sup>+</sup>* mice with MPN and *Ptpn11<sup>+/+</sup>/Nestin-Cre<sup>+</sup>* littermates ( $n = 4$  mice per group) were assayed for haematopoietic colony-forming units in 0.9% methylcellulose IMDM medium containing 30% FBS, glutamine ( $10^{-4}$  M),  $\beta$ -mercaptoethanol ( $3.3 \times 10^{-5}$  M), and IL-3 ( $1.0$  ng ml $^{-1}$ ) or GM-CSF ( $1.0$  ng ml $^{-1}$ ). After 7 days of culture at 37°C in a humidified 5% CO $_2$  incubator, haematopoietic cell colonies (primarily CFU-GM) derived from myeloid progenitors were counted under an inverted microscope. **c**, Femurs, spleens, livers and lungs were processed for histopathological examination (haematoxylin and eosin staining) ( $n = 4$  mice per group). Representative pictures are shown. **d–h**, BM cells and splenocytes were collected from *Ptpn11<sup>E76K/+</sup>/Nestin-Cre<sup>+</sup>*

mice with MPN and *Ptpn11<sup>+/+</sup>/Nestin-Cre<sup>+</sup>* littermates. CD115<sup>+</sup>Gr-1<sup>+</sup> monocytes in the BM ( $n = 4$  mice per group) (**d**), frequencies of LSKs (Lin<sup>-</sup>Sca-1<sup>+</sup>c-Kit<sup>+</sup>) in the BM ( $n = 8$  mice per group) (**e**), absolute number of HSCs in two femurs and two tibias ( $n = 10$  mice per group) (**f**), frequencies of LSKs in the spleen ( $n = 8$  mice per group) (**g**), and apoptotic cells in the HSC population in the BM ( $n = 6$  mice per group) (**h**) were assayed by multiparameter FACS analyses. **i**, BM cells collected from wild-type Boyl mice were transplanted into 6-month-old *Ptpn11<sup>E76K/+</sup>/Nestin-Cre<sup>+</sup>* and *Ptpn11<sup>+/+</sup>/Nestin-Cre<sup>+</sup>* mice. Recipients were monitored for MPN development for 6–8 months. Percentages of donor cell (CD45.1<sup>+</sup>)-derived Mac-1<sup>+</sup> myeloid cells in the peripheral blood ( $n = 5$  mice per group) and BM ( $n = 8$  mice per group) of recipients were determined. **j**, Frequencies of HSCs in the BM from *Ptpn11<sup>E76K/+</sup>/Nestin-Cre<sup>+</sup>* mice that had not yet manifested MPN and *Ptpn11<sup>+/+</sup>/Nestin-Cre<sup>+</sup>* littermates ( $n = 4$  mice per group) were assayed as above. Data shown in **a**, **b**, **d–j** are mean  $\pm$  s.d. of all mice examined; \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . Source Data for this figure are available online.









**Extended Data Figure 4 | *Ptpn11*<sup>E76K/+</sup> mutation in MSCs and osteoprogenitors, but not differentiated osteoblasts or endothelial cells, in the BM microenvironment induces MPN.** Cell-type-specific *Ptpn11*<sup>E76K</sup> knock-in mice as indicated were generated and monitored for MPN development. **a**, The ages of the microenvironmental cell-type-specific *Ptpn11*<sup>E76K/+</sup> knock-in mice when they were euthanized for MPN diagnosis. **b**, Peripheral blood haematology was determined using the HemaTrue veterinary hematology analyzer. Mac-1<sup>+</sup>Gr-1<sup>+</sup> myeloid cells, B220<sup>+</sup> B lymphoid, and CD3<sup>+</sup> T-lymphoid cells in the BM were analysed by FACS. Karyotypes of MPN cells were examined by standard karyotyping analyses. HSCs in the peripheral blood were determined by multiparameter FACS. SSC/CD45 profiles were also determined by FACS.

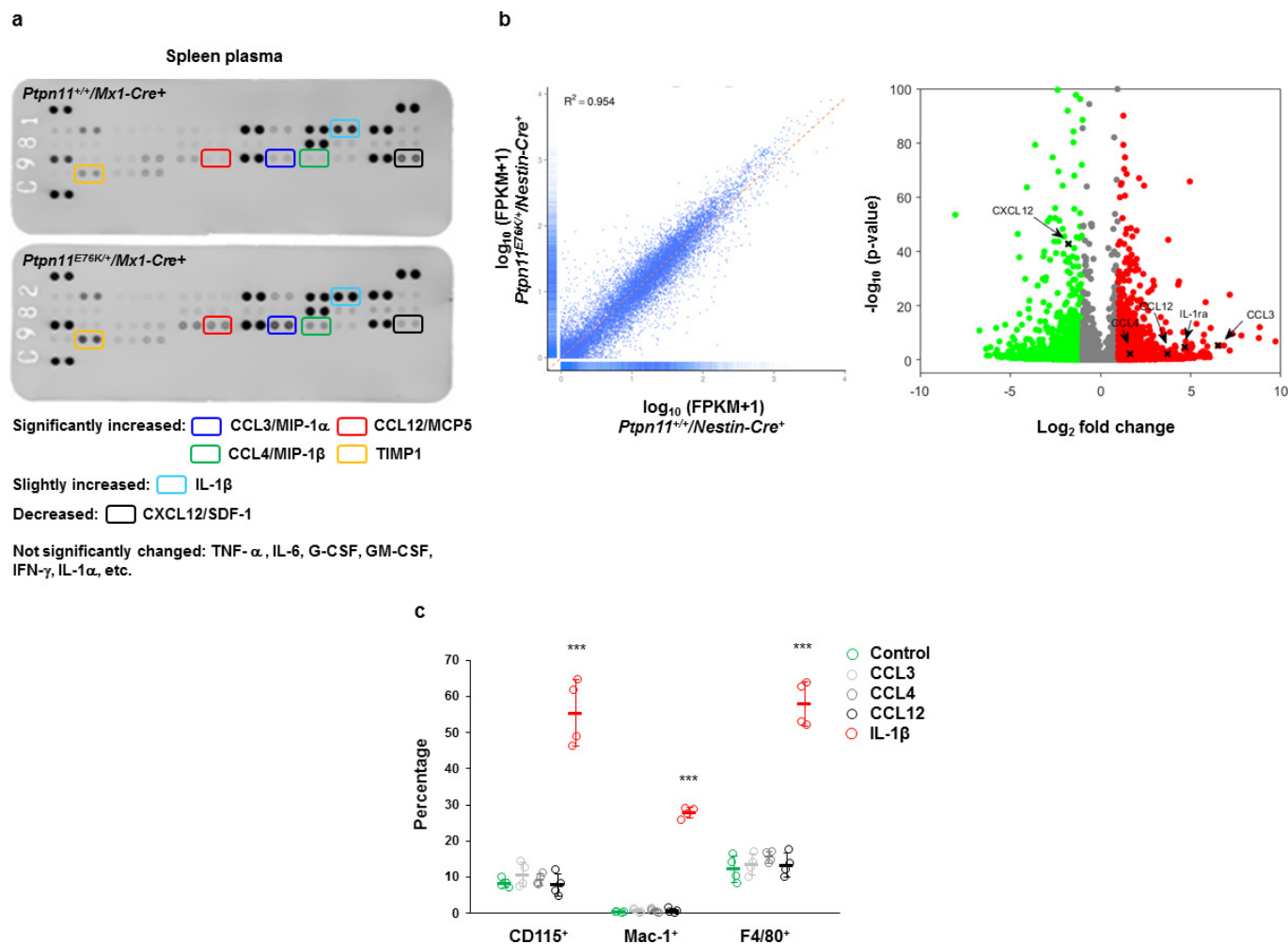
CD115<sup>+</sup>Gr-1<sup>+</sup> monocytes were highlighted in red. Cytokine sensitivity of BM myeloid progenitors was determined by CFU assays with a range of GM-CSF concentrations. Transplantability of MPN cells was determined by transplantation of BM cells into lethally-irradiated BoyJ mice. Recipient mice were monitored for 6 months. All methods are described in Methods and/or related figure legends. **c**, BM-derived MSCs were generated from the indicated mouse lines. The abundance of the neo cassette in genomic DNA was determined by qPCR ( $n = 3$  mice per group). Data shown in **c** are mean  $\pm$  s.d. of all mice examined. Statistical significance (\*\*\*)  $P < 0.001$  was determined between the indicated cell-type-specific *Ptpn11*<sup>E76K/+</sup> knock-in mice and *Ptpn11*<sup>E76K-neo/+</sup> *Nestin-Cre*<sup>-</sup> control mice. Source Data for this figure are available online.





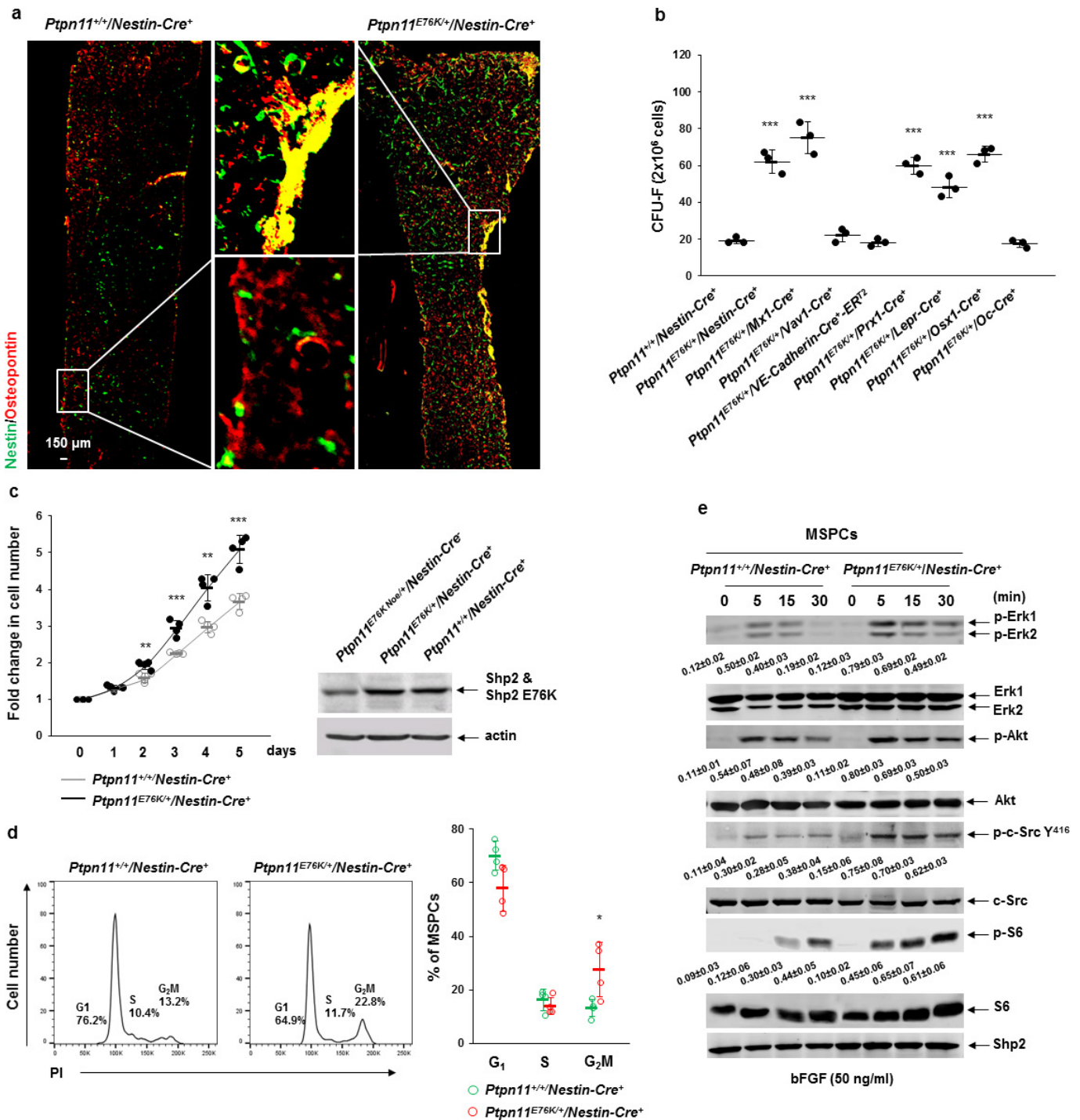
**Extended Data Figure 5 | *Ptpn11*<sup>E76K/+</sup> MSPCs do not directly activate HSCs.** BM-derived MSPCs were enriched from *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>*Nestin-Cre*<sup>+</sup> mice, as described in Methods. MSPCs at the 2nd or 3rd passages were plated in regular 24-well plates (a) or lower chambers of transwells (b). Forty-eight hours later when the cells were confluent, HSCs (75–200) (Lin<sup>−</sup>Sca-1<sup>+</sup>c-Kit<sup>+</sup>CD150<sup>+</sup>CD48<sup>−</sup>Flk2<sup>−</sup>) sorted from *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>*Mx1-Cre*<sup>+</sup> mice (8 weeks after pI-pC administration) were seeded in the same wells (a) or in upper chambers with the 0.4 μm pore size (b). The cells were

co-cultured in StemSpan medium supplemented with cytokines TPO (50 ng ml<sup>−1</sup>), Flt3 ligand (50 ng ml<sup>−1</sup>), SCF (50 ng ml<sup>−1</sup>), IL-3 (20 ng ml<sup>−1</sup>), and IL-6 (20 ng ml<sup>−1</sup>). Frequencies of myeloid (Mac-1<sup>+</sup>Gr-1<sup>+</sup>) cells that differentiated from HSCs were assayed by FACS analyses after 7–10 days of co-culture. Experiments were performed three times and similar results were obtained in each (see Supplementary Information). Results shown are mean ± s.d. of triplicates from one experiment; N.S., not significant. Source data are available online.



**Extended Data Figure 6 | Cytokines/chemokines are aberrantly produced by *Ptpn11*<sup>E76K/+</sup> MSPCs.** **a**, Spleen tissues freshly dissected from *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>*Mx1-Cre*<sup>+</sup> mice ( $n = 3$  mice per group) 12 weeks after pI-pC administration were gently smashed in PBS (0.1 g tissue per 1.0 ml). Supernatant collected was processed for cytokine-chemokine array analyses with the Mouse Cytokine Antibody Array Kit following the instructions provided by the manufacturer. Representative results from one pair of the mice are shown. **b**, MSPCs (CD45<sup>+</sup>Ter-119<sup>+</sup>CD31<sup>+</sup>CD140α<sup>+</sup>Sca-1<sup>+</sup>) were freshly isolated from paired *Ptpn11*<sup>E76K/+</sup>*Nestin-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>*Nestin-Cre*<sup>+</sup> mice at 7–8 months old by FACS. Total RNA was extracted and processed for

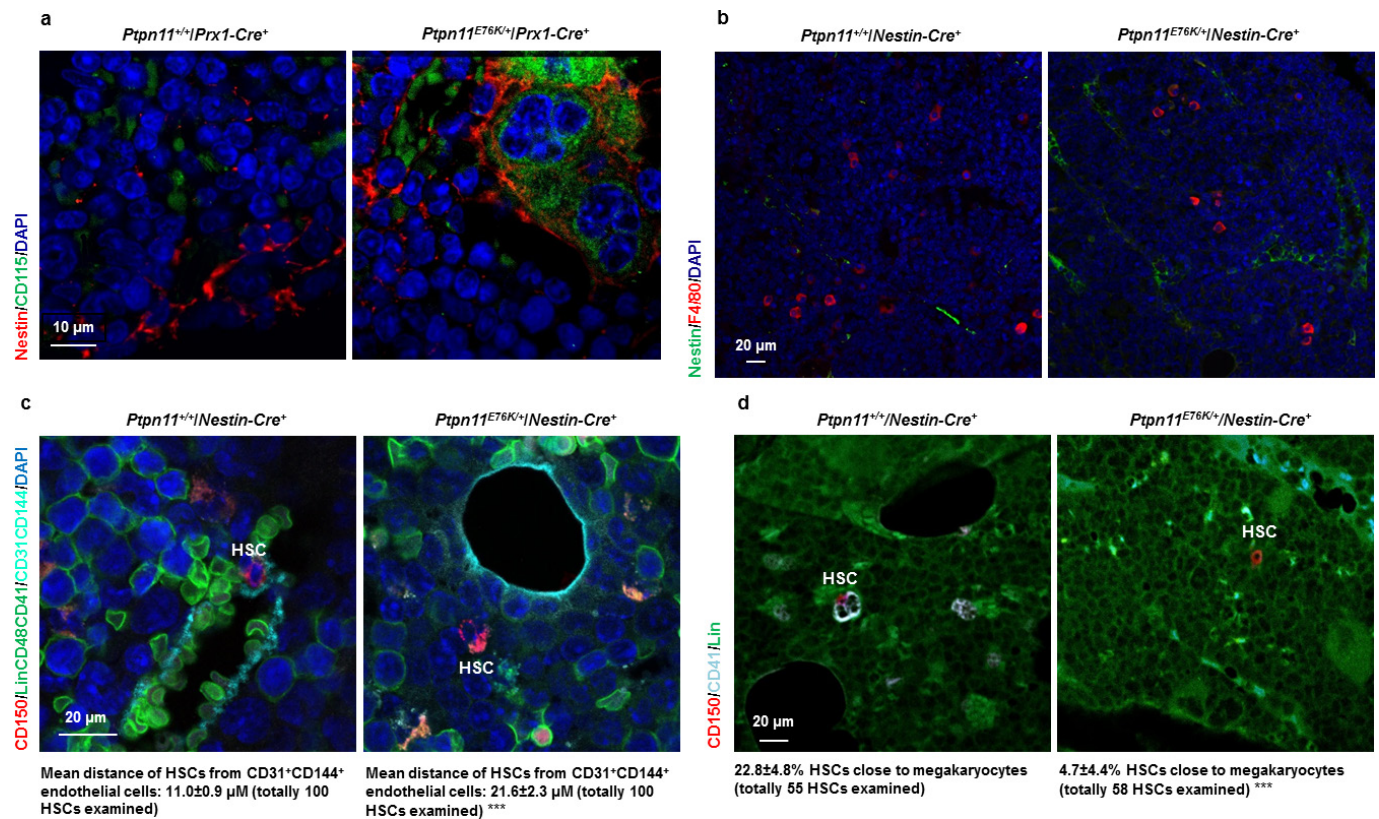
RNA-sequencing analyses as described in Methods. The correlation coefficient between the two groups was 0.954, verifying that the method was accurate (left). Genes with more than 2.0 fold increased (in red) or decreased (in green) mRNA levels are shown on the right. Secreted protein factors are indicated. **c**, HSCs sorted from wild-type C57BL/6 mice were cultured in the presence of IL-1β (10 ng ml<sup>-1</sup>), CCL3 (20 ng ml<sup>-1</sup>), CCL4 (20 ng ml<sup>-1</sup>), or CCL12 (20 ng ml<sup>-1</sup>). Six days later, cells were collected and analysed for Mac-1<sup>+</sup> myeloid cells, F4/80<sup>+</sup> macrophages, and CD115<sup>+</sup> monocytes by FACS. Data presented are mean ± s.d. of four independent experiments; \*\*\* $P < 0.001$ . Source Data for this figure are available online.



**Extended Data Figure 7 | *Ptpn11*<sup>E76K/+</sup> mutation increases MSPC proliferation by enhancing cell signalling activities.** **a**, Seven-month-old *Ptpn11*<sup>E76K/+</sup>/*Nestin-Cre*<sup>+</sup> mice and *Ptpn11*<sup>+/+</sup>/*Nestin-Cre*<sup>+</sup> littermates were analysed. Femurs were processed for immunofluorescence staining with the indicated antibodies ( $n = 3$  mice per group). Representative images are shown. **b**, BM cells ( $2 \times 10^6$  cells) freshly collected from the indicated mouse lines ( $n = 3$  mice per group) were assessed by the CFU-F assay, as detailed in Methods. Statistical significance was determined between the indicated cell-type-specific *Ptpn11*<sup>E76K/+</sup> knock-in mice and *Ptpn11*<sup>+/+</sup>/*Nestin-Cre*<sup>+</sup> control mice. **c**, **d**, MSPCs were enriched from *Ptpn11*<sup>E76K/+</sup>/*Nestin-Cre*<sup>+</sup> and *Ptpn11*<sup>+/+</sup>/*Nestin-Cre*<sup>+</sup> mice ( $n = 4$  mice per group) as described in Methods. MSPCs were analysed for growth rates (**c**, left) and expression levels of Shp2 (**c**, right), and cell cycle distributions (**d**). **e**, Confluent

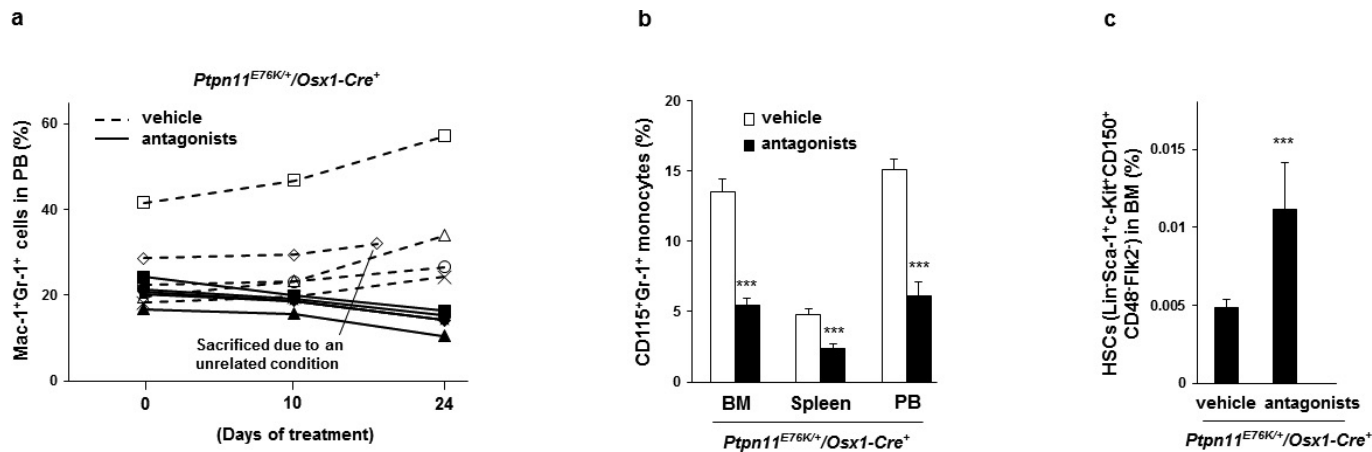
MSPCs ( $n = 3$  mice per group) were starved in serum and growth factor-free medium for 48 h and then stimulated with basic fibroblast growth factor (bFGF,  $50 \text{ ng ml}^{-1}$ ) for the indicated periods of time. Whole-cell lysates were prepared and examined for Erk, Akt, c-Src, and S6 activities by immunoblotting with anti-phospho-Erk, anti-phospho-Akt, anti-phospho-c-Src Y<sup>416</sup>, and anti-phospho-S6 antibodies. Blots were stripped and reprobed with anti-pan-Erk, anti-pan-Akt, anti-c-Src, anti-S6, and anti-Shp2 antibodies to check protein loading and Shp2 levels. Densitometric analyses were performed to determine phosphorylation levels of the indicated proteins and normalized against protein loading levels (arbitrary units). Data shown in **b–e** are mean  $\pm$  s.d. of all mice examined. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . Source Data for this figure are available online.





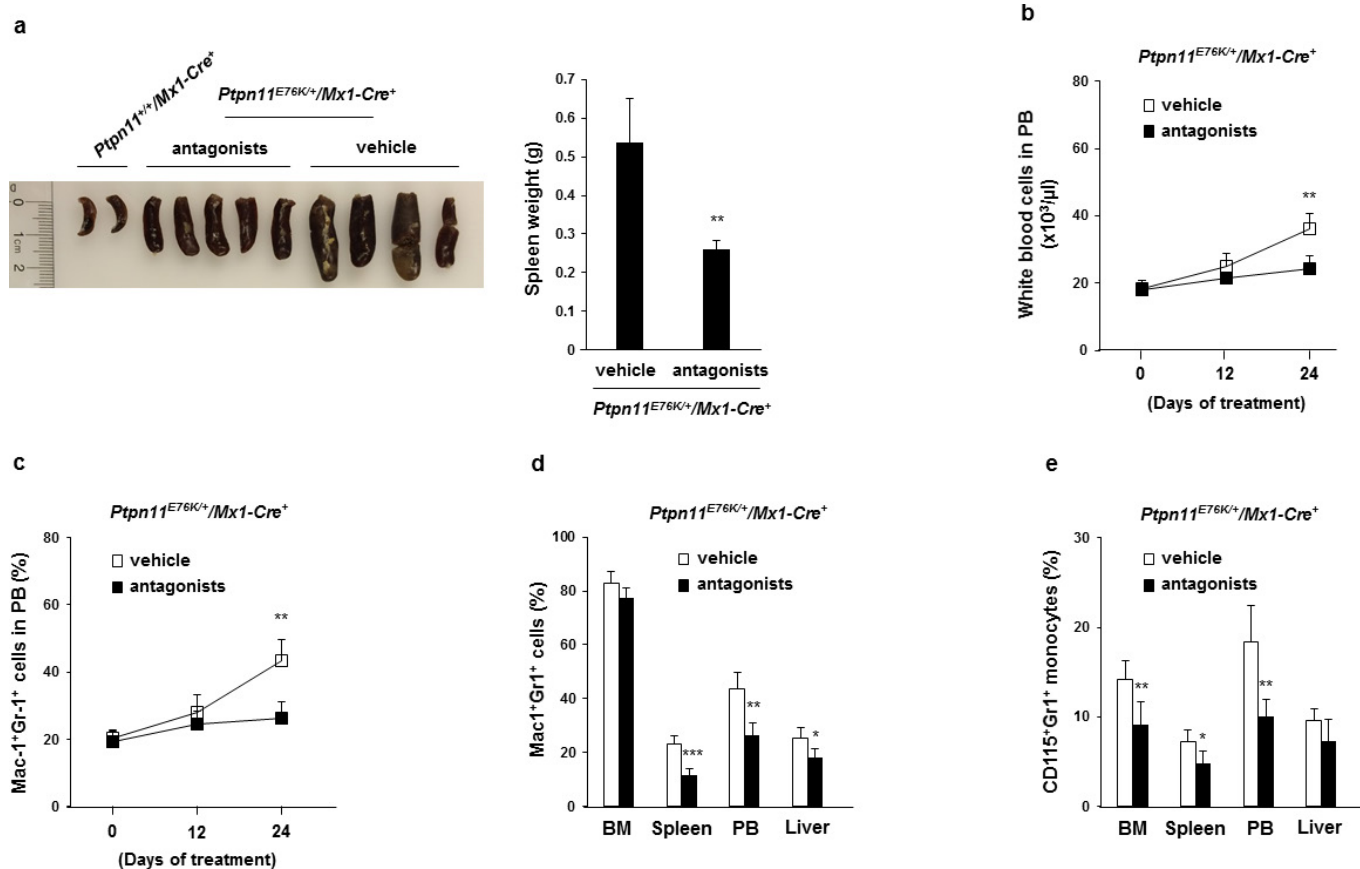
**Extended Data Figure 8 | *Ptpn11<sup>E76K/+</sup>* MSPCs recruit monocytes, but not macrophages, to the vicinity. **a**, Bone sections (one section per femur or tibia) prepared from *Ptpn11<sup>E76K/+</sup>/Prx1-Cre<sup>+</sup>* ( $n=4$  mice) and *Ptpn11<sup>+/+</sup>/Prx1-Cre<sup>+</sup>* mice ( $n=3$  mice) at 6–7 months old were immunostained with the indicated antibodies and counterstained with DAPI. **b–d**, Bone sections (one section per femur or tibia) prepared from *Ptpn11<sup>E76K/+</sup>/Nestin-Cre<sup>+</sup>* and *Ptpn11<sup>+/+</sup>/Nestin-Cre<sup>+</sup>* mice at 7–10 months old were immunostained with the indicated antibodies and**

counterstained with DAPI ( $n=5$  mice per group) (**b**). The distance of HSCs (Lin<sup>−</sup>CD48<sup>−</sup>CD41<sup>−</sup>CD150<sup>+</sup>) from closest CD31<sup>+</sup>CD144<sup>+</sup> endothelial cells was determined ( $n=8$  mice per group) (**c**). The spatial relationship between HSCs (Lin<sup>−</sup>CD41<sup>−</sup>CD150<sup>+</sup>) and megakaryocytes (CD41<sup>+</sup>) was examined. HSCs within  $<8\mu\text{m}$  of megakaryocytes were considered as close to megakaryocytes ( $n=5$  mice per group) (**d**). Representative images are shown in all panels.



**Extended Data Figure 9 | Administration of CCL3 receptor antagonists reverses MPN phenotypes in *Ptpn11<sup>E76K/+</sup>Osx1-Cre<sup>+</sup>* mice.** *Ptpn11<sup>E76K/+</sup>Osx1-Cre<sup>+</sup>* mice at 6–7 months old were treated daily with the CCR1 antagonist BX471 and the CCR5 antagonist Maraviroc or vehicle control for 23 days as described in Methods. Myeloid cells (Mac-1<sup>+</sup>Gr-1<sup>+</sup>) in the peripheral blood were determined at the indicated time points ( $n = 5$  mice per group, each line represents one mouse) (a). Mice were euthanized at the end of the experiments. Monocytes

(CD115<sup>+</sup>Gr-1<sup>+</sup>) ( $n = 5$  and 6 mice for the antagonist and vehicle groups, respectively) in the BM, spleen, and peripheral blood were determined (b). BM cells were assayed by multiparameter FACS analyses to determine the pool size (c) of HSCs (Lin<sup>-</sup>Sca-1<sup>+</sup>c-Kit<sup>+</sup>CD150<sup>+</sup>CD48<sup>-</sup>Flk2<sup>-</sup>) ( $n = 5$  and 6 mice for the antagonist and vehicle groups, respectively). Data shown in b, c are mean  $\pm$  s.d. of all mice examined; \*\*\* $P < 0.001$ . Source Data for this figure are available online.



**Extended Data Figure 10 | Administration of CCL3 receptor antagonists mitigates MPN in *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> mice.** *Ptpn11*<sup>E76K/+</sup>*Mx1-Cre*<sup>+</sup> mice (4 weeks after pI-pC administration; *n* = 5 mice per group) were treated daily with the CCR1 and CCR5 antagonists or vehicle as described above. Mice were euthanized, and spleens were photographed and weighted (a). White blood cell counts (b) and myeloid

cells (Mac-1<sup>+</sup>Gr-1<sup>+</sup>) (c) in the peripheral blood were determined at the indicated time points. Mac-1<sup>+</sup>Gr-1<sup>+</sup> myeloid cells (d) and CD115<sup>+</sup>Gr-1<sup>+</sup> monocytes (e) in the BM, spleen, peripheral blood, and liver were determined at the end of the experiments. Data shown in all panels are mean ± s.d. of all mice examined; \**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.001. Source Data for this figure are available online.



# Single-cell RNA-seq supports a developmental hierarchy in human oligodendrogloma

Itay Tirosh<sup>1\*</sup>, Andrew S. Venteicher<sup>1,2,3\*</sup>, Christine Hebert<sup>1,2</sup>, Leah E. Escalante<sup>1,2</sup>, Anoop P. Patel<sup>3</sup>, Keren Yizhak<sup>1,2</sup>, Jonathan M. Fisher<sup>1</sup>, Christopher Rodman<sup>1</sup>, Christopher Mount<sup>4</sup>, Mariella G. Filbin<sup>1,2,5</sup>, Cyril Neftel<sup>1,2</sup>, Niyati Desai<sup>2</sup>, Jackson Nyman<sup>1</sup>, Benjamin Izar<sup>1</sup>, Christina C. Luo<sup>2</sup>, Joshua M. Francis<sup>1,6</sup>, Aanand A. Patel<sup>2</sup>, Maristela L. Onozato<sup>2</sup>, Nicolo Riggi<sup>2</sup>, Kenneth J. Livak<sup>1</sup>, Dave Gennert<sup>1</sup>, Rahul Satija<sup>1</sup>, Brian V. Nahed<sup>3</sup>, William T. Curry<sup>3</sup>, Robert L. Martuza<sup>3</sup>, Ravindra Mylvaganam<sup>2</sup>, A. John Iafrate<sup>2</sup>, Matthew P. Frosch<sup>2</sup>, Todd R. Golub<sup>1,5,7</sup>, Miguel N. Rivera<sup>1,2</sup>, Gad Getz<sup>1,2</sup>, Orit Rozenblatt-Rosen<sup>1</sup>, Daniel P. Cahill<sup>3</sup>, Michelle Monje<sup>4</sup>, Bradley E. Bernstein<sup>1,2</sup>, David N. Louis<sup>2</sup>, Aviv Regev<sup>1,7§</sup> & Mario L. Suvà<sup>1,2§</sup>

**Although human tumours are shaped by the genetic evolution of cancer cells, evidence also suggests that they display hierarchies related to developmental pathways and epigenetic programs in which cancer stem cells (CSCs) can drive tumour growth and give rise to differentiated progeny<sup>1</sup>. Yet, unbiased evidence for CSCs in solid human malignancies remains elusive. Here we profile 4,347 single cells from six *IDH1* or *IDH2* mutant human oligodendrogliomas by RNA sequencing (RNA-seq) and reconstruct their developmental programs from genome-wide expression signatures. We infer that most cancer cells are differentiated along two specialized glial programs, whereas a rare subpopulation of cells is undifferentiated and associated with a neural stem cell expression program. Cells with expression signatures for proliferation are highly enriched in this rare subpopulation, consistent with a model in which CSCs are primarily responsible for fuelling the growth of oligodendrogloma in humans. Analysis of copy number variation (CNV) shows that distinct CNV sub-clones within tumours display similar cellular hierarchies, suggesting that the architecture of oligodendrogloma is primarily dictated by developmental programs. Subclonal point mutation analysis supports a similar model, although a full phylogenetic tree would be required to definitively determine the effect of genetic evolution on the inferred hierarchies. Our single-cell analyses provide insight into the cellular architecture of oligodendrogliomas at single-cell resolution and support the cancer stem cell model, with substantial implications for disease management.**

Intra-tumoural heterogeneity contributes to therapy failure and cancer progression<sup>1</sup>. Although branched genetic evolution of cancer cells is a key determinant of tumour heterogeneity, non-genetic programs such as those associated with the self-renewal of tissue stem cells and their differentiation into specialized cell types contribute further to tumour functional heterogeneity. In human gliomas, candidate CSCs have been functionally isolated in high-grade (WHO grade III–IV) lesions<sup>2</sup>. However, functional approaches such as *in vivo* orthotopic xenotransplantation in mice or *in vitro* sphere formation assays have generated controversy, as they identify candidate CSCs through selection in xenogeneic environments that are very different from the native tumour milieu and only provide limited genetic characterization of putative CSCs. In addition, it remains unknown if gliomas contain CSCs early in their development—as grade II lesions—a question central to our understanding of the initial steps of gliomagenesis<sup>3</sup>. Thus, it is critical to develop a framework that allows the analysis of cellular

programs at single-cell resolution and across different genetic clones in human tumours *in situ* at each stage of clinical progression.

We focused on oligodendrogloma, an incurable glioma characterized by mutations in *IDH1* or *IDH2* and co-deletion of chromosome arms 1p and 19q<sup>4</sup>. We performed single-cell RNA-seq<sup>5</sup> (scRNA-seq) from six untreated grade II oligodendrogliomas, in which *IDH1* or *IDH2* mutation and 1p/19q co-deletion were confirmed (Extended Data Fig. 1a–c). Overall, we analysed 4,347 cells that passed quality controls (Methods; Extended Data Fig. 1d). Three tumours were analysed more deeply (MGH36, MGH53 and MGH54, with analysis of 791 to 1,229 cells per tumour) and three (MGH60, MGH93 and MGH97) were profiled at medium depth (430 to 598 cells analysed).

We distinguished malignant from non-malignant cells by estimating CNV from the average expression of genes in large chromosomal regions within each cell<sup>6</sup> (Fig. 1a; Methods). Each tumour contained a large majority of cells with the 1p/19q co-deletion, as well as some cases of tumour-specific CNVs, which were validated by fluorescence *in situ* hybridization (FISH) and by whole-exome sequencing (WES) (Fig. 1a and Extended Data Fig. 1c). In two tumours (MGH36 and MGH97), CNV analysis identified two sub-clones (Fig. 1a, b).

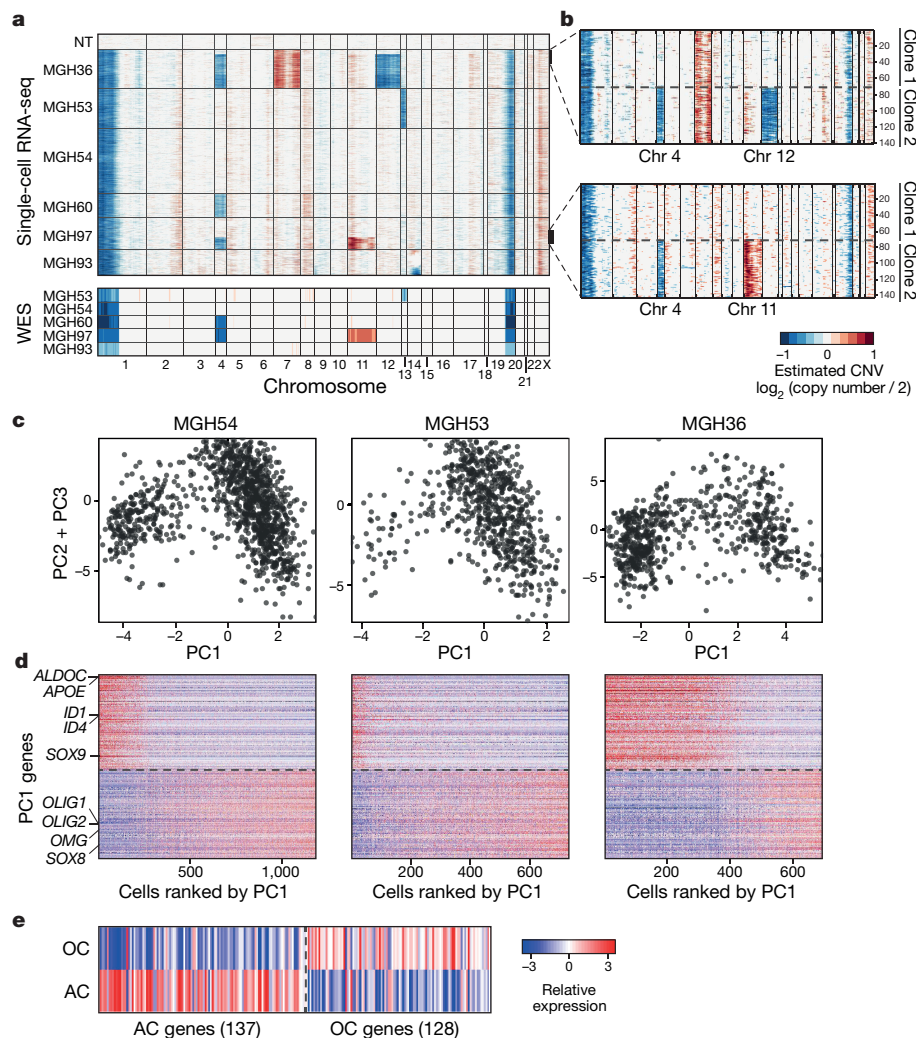
Another 303 cells lacked detectable CNVs and clustered by gene expression into subsets that expressed microglia and mature oligodendrocyte markers, respectively (Extended Data Fig. 2a). There was significant variation between the microglia cells, with a set of pro-inflammatory cytokines (IL1A, IL1B, IL8 and TNF), chemokines (CCL3, CCL4) and early response genes coordinately expressed by ~80% of the microglia (Extended Data Fig. 2b). This program differs from canonical macrophage M1 and M2 responses<sup>7</sup>, suggesting an unknown microglia program that may be glioma-specific.

We next examined cancer cell heterogeneity from the three tumours with the largest cell numbers. A combined principal component analysis (PCA) (Methods) identified two prominent groups of cells, corresponding to low and high PC1 scores (Fig. 1c) and expressing distinct lineage markers of astrocytes and oligodendrocytes, respectively. These results were highly consistent across all six tumours, and were not accounted for by technical or batch effects (Extended Data Fig. 2c–f and Supplementary Note 1). In each tumour, cells with high PC1 scores were strongly associated with the high expression of 137 genes, including oligodendrocytic markers (for example, *OLIG1*, *OLIG2*, *OMG*), and with the low expression of 128 genes, including astrocytic markers (for example, *APOE*, *ALDOC*, *SOX9*) (Fig. 1d, e and

<sup>1</sup>Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA. <sup>2</sup>Department of Pathology and Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA. <sup>3</sup>Department of Neurosurgery, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA. <sup>4</sup>Departments of Neurology, Neurosurgery, Pediatrics and Pathology, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>5</sup>Department of Pediatric Oncology, Dana-Farber Cancer Institute and Children's Hospital Cancer Center, Boston, Massachusetts 02215, USA. <sup>6</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. <sup>7</sup>Howard Hughes Medical Institute, Koch Institute, Department of Biology, MIT, Cambridge, Massachusetts 02139, USA.

\*These authors contributed equally to this work.

§These authors jointly supervised this work.



**Figure 1 | Single-cell RNA-seq of cancer and non-cancer cells in six oligodendrogliomas.** **a**, CNV profiles inferred from scRNA-seq (top) and DNA whole-exome sequencing (WES) (bottom) of oligodendrogliomas. Cells (rows,  $n = 4,347$ ) are ordered from non-tumoural cells (NT,  $n = 303$ ) to cancer cells ( $n = 4,044$ ), ordered into six oligodendrogliomas. **b**, In MGH36 and MGH97, cells are ordered by CNVs, with zoomed-in view shown. **c**, PCA of malignant cells. Shown are PC1 (x axis) versus PC2 and PC3 (y axis) scores of cells from three tumours based on a single combined

Supplementary Table 1)<sup>8,9</sup>. Cells with low PC1 scores had the opposite patterns of expression. This suggests that oligodendrogliomas are primarily composed of two subpopulations of glial cells, and this mirrors the histopathology<sup>4</sup>.

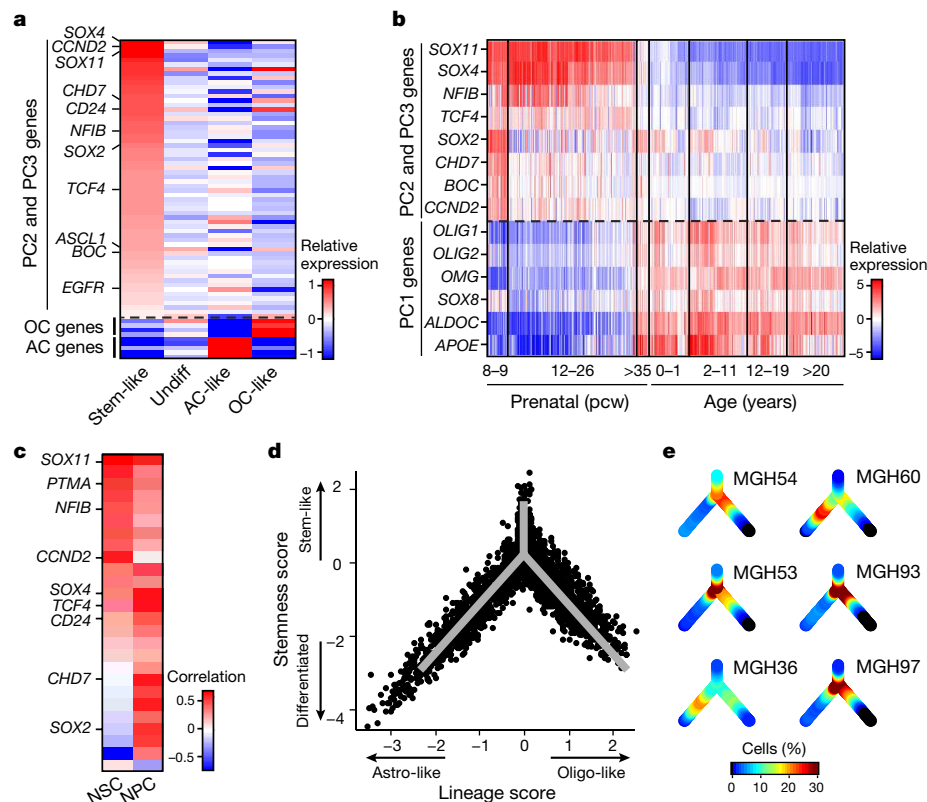
Cells with high PC2 and PC3 scores had intermediate PC1 scores (Fig. 1c and Extended Data Fig. 2c, e), suggesting a lack of differentiation, and prompting us to explore additional programs. A total of 63 genes were associated with high PC2 and PC3 (Fig. 2a, Supplementary Table 1 and Methods), and several lines of evidence suggest that they represent a 'stemness' program. The 20 highest-ranking genes include *SOX4*, *SOX11* and *SOX2*, neurodevelopmental transcription factors critical to neural stem cells and glioma CSCs<sup>10–12</sup>. Additional signature genes with important roles in neurogenesis and in glioma CSCs included *NFIB*, *ASCL1*, *CHD7*, *CD24*, *BOC* and *TCF4* (refs 6, 10–14). Similar results were obtained by hierarchical clustering, which showed a distinct cluster of cells preferentially expressing PC2- and PC3-associated genes (Extended Data Fig. 3a, b). Several of these genes were identified by scRNA-seq in primary glioblastoma CSCs (Extended Data Fig. 3c,  $P = 1.5 \times 10^{-4}$ , hypergeometric test). Expression of PC2- and PC3-associated regulators was highest in prenatal human brain and dropped significantly after birth, suggesting a role in early neural development

PCA. **d**, Astrocyte-like and oligodendrocyte-like signatures. Relative expression of genes correlated most positively (bottom) or negatively (top) with PC1, in cancer cells from each of the three tumours (marked as in c), ranked by PC1 scores. Selected astrocyte (AC) and oligodendrocyte (OC) marker genes are highlighted. **e**, Relative expression of the mice orthologues of the genes shown in **d** ( $\log_2$ -ratio of the respective cell type compared to the average of oligodendrocyte, astrocyte, OPC and neurons)<sup>9</sup>.

(Allen Brain Atlas<sup>15</sup>, Fig. 2b,  $P = 8 \times 10^{-18}$ ,  $t$ -test). Similarly, PC2- and PC3-associated genes were preferentially expressed in single cells from fetal human brain ( $P = 0.006$ , hypergeometric test)<sup>16</sup>. On the basis of these lines of evidence, we separated cells with intermediate PC1 values into 'undifferentiated' (low PC2 and PC3) and 'stem/progenitors' (high PC2 and PC3) cells (Fig. 2a).

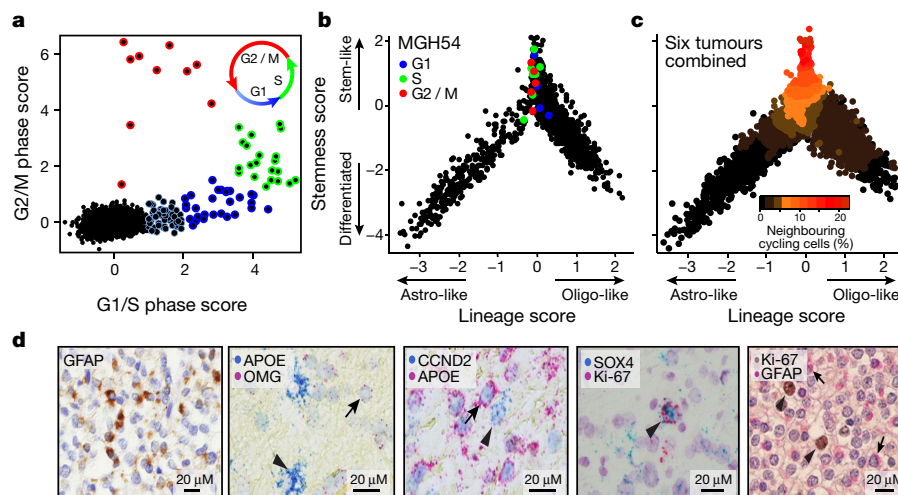
Oligodendrogliomas are thought to arise from transformation of oligodendrocyte progenitor cells (OPCs)<sup>17</sup>. However, PC2 and PC3 genes were not preferentially expressed in OPCs; instead, they were preferentially expressed in cells of neuronal lineage<sup>9,18</sup> (Extended Data Fig. 3d) and upregulated upon activation of tri-potent mouse neural stem cells<sup>19</sup> (NSCs) (Fig. 2c, Extended Data Fig. 3e;  $P = 3 \times 10^{-6}$ ,  $t$ -test).

To further examine if the stemness program is associated with tri-potent stem/progenitor cells, we profiled human neural progenitor cells (NPCs) by scRNA-seq (Extended Data Fig. 4a–d). PCA of the NPC profiles identified an expression program highly similar to the PC2 and PC3 associated program of tumour cells (Fig. 2c, Extended Data Figs 3f and 4e, f and Supplementary Table 2;  $P = 2 \times 10^{-35}$ ,  $t$ -test). Thus, a common program is shared by subsets of putative oligodendrogloma stem cells and normal NPCs and NSCs. Together, our analysis reveals three main expression patterns recapitulating oligodendrocytic and



**Figure 2 | Stemness expression program and a developmental hierarchy in oligodendroglioma.** **a**, Average relative expression of genes most highly correlated with PC2 and PC3 (top), and selected astrocyte and oligodendrocyte genes (bottom), in stem-like cells, undifferentiated cells, oligodendrocyte-like and astrocyte-like cells (Methods). Genes were sorted by relative expression in stem-like cells. **b**, Stemness genes are preferentially expressed in early human brain development. Relative expression of PC2 and PC3 putative stemness genes (top) and oligodendrocyte and astrocyte marker genes (bottom) across 524 human brain samples (Allen Brain Atlas). Samples are ordered in columns by age, from prenatal (left) to adult (right). **c**, The stemness program is

correlated to mouse activated NSC and human NPCs. Pearson correlation coefficients between the expression of PC2 and PC3 genes (rows) and expression programs of mouse NSC activation<sup>19</sup> (left) and human NPCs (right) across single cells from the respective datasets (Extended Data Figs 3e, f and 4). **d**, Inferred developmental hierarchy in oligodendroglioma cells ( $n = 4,044$ ). Lineage and stemness scores (Methods) of malignant cells from the six tumours. Grey lines indicate the 'backbone' used in **e** and Extended Data Fig. 8b. **e**, Colour-coded density of cells (fraction of cells within a Euclidean distance of 0.3) from each tumour across the backbone of the hierarchy.



**Figure 3 | Cycling cells are enriched among oligodendroglioma stem/progenitor cells.** **a**, Classification of cells ( $n = 4,044$ ) to non-cycling (black) and categories of cycling cells (colour-coded by approximated phase as per inset) based on the relative expression of gene-sets associated with G1/S ( $x$  axis) and G2/M ( $y$  axis). **b**, **c**, Only stem/progenitor cells are cycling. **b**, Hierarchy plot, as in Fig. 2d, for MGH54 cells ( $n = 1,174$ ), with cycling cells colour-coded as in **a**. **c**, Hierarchy plot for the six tumours, with each cell colour-coded based on the fraction of neighbouring cells (within Euclidean distance of 0.3) that are cycling.

**d**, Immunohistochemistry for astrocytic marker (GFAP) in MGH54, with expression in subset of cells (left). *In situ* RNA hybridization shows mutually exclusive expression of astrocytic (APOE, arrowhead) and oligodendrocytic (OMG, arrow) markers, and of stem/progenitor (CCND2, arrowhead) and APOE (arrow) markers, but co-expression of stemness (SOX4) and cell cycle (Ki-67) markers (arrowhead) (middle). Double immunohistochemistry for GFAP (red, arrows) and Ki-67 (brown, arrowheads), showing mutual exclusivity (right).



astrocytic differentiation, and stem/progenitor programs of early neural development.

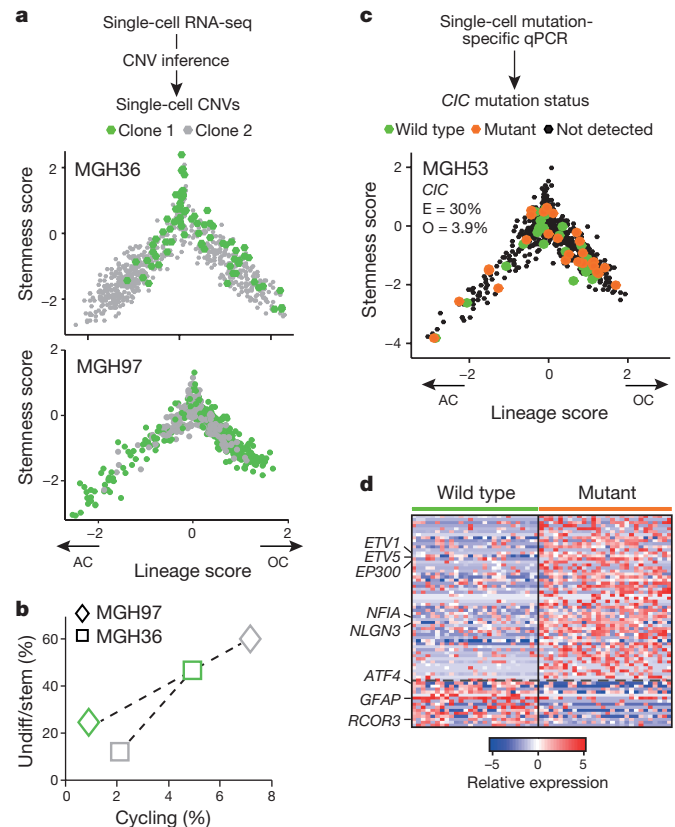
To assign a cellular state to each tumour cell precisely, we defined differentiation and stemness scores (Methods). Plotting these scores across the cells of all six tumours revealed similarity to normal development (Fig. 2d), with a transition from stem/progenitor programs into differentiation along two glial lineages. Notably, the same architecture was observed in each of the six tumours and also found when tested with a different method for scRNA-seq (Fig. 2e, Extended Data Fig. 5a, e, Methods). Statistical analysis of the lineage scores suggests the existence of intermediate states for each lineage (Extended Data Fig. 5c and Supplementary Note 2).

To assess how tumour cell proliferation and self-renewal may relate to developmental programs, we next scored each cell for the expression of signatures for the G1/S and G2/M phases (Methods)<sup>20,21</sup>. We found a small proportion of cells in each tumour (1.5–8%) that were proliferating, consistent with Ki-67 staining, and we estimated the cell-cycle phase of proliferating cells (Fig. 3a, Extended Data Fig. 6a–c and Supplementary Table 3). Almost all proliferating cancer cells were confined to the stem/progenitor and undifferentiated subpopulation of the tumour (Fig. 3b, c, Extended Data Fig. 6d and Supplementary Table 3), suggesting that this is the compartment fuelling the growth of oligodendroglioma in humans. We confirmed these patterns in tumours by both RNA *in situ* hybridization and immunohistochemistry with markers of astrocytes (GFAP and APOE), oligodendrocytes (OMG), stem/progenitor cells (SOX4, CCND2) and cell proliferation (Ki-67) in tissue staining across the original six tumours and in a validation cohort of ten additional tumours (Fig. 3d, Extended Data Figs 5d and 8c and Supplementary Table 3). Additionally, there is a strong correlation between our cell-cycle and stem/progenitor signatures across 69 bulk oligodendroglioma samples in The Cancer Genome Atlas<sup>22</sup> (Extended Data Fig. 6e). Finally, the enrichment of cell-cycle signatures among stem/progenitor and undifferentiated cells was even more striking for cells inferred to be in G2/M phases compared to those in G1 phase (Extended Data Fig. 6f), possibly reflecting a short G1 phase in stem cells<sup>23</sup>.

Although cycling cells were highly enriched among stem/progenitors, their frequency was low (~10%) even in that compartment; accordingly, the PC2 and PC3 program did not include a signature for the cell cycle, except for *CCND2* (Fig. 2a), a gene controlling cell cycle and self-renewal of glioma CSCs<sup>24</sup>. *CCND2* was highly expressed both in cycling cells and in non-cycling stem/progenitor cells (Extended Data Fig. 7a, b), consistent with *CCND2* priming cells for the cell cycle<sup>23</sup>. Interestingly, stem/progenitor tumour cells preferentially express *CCND2*, whereas differentiated cells express *CCND1* and *CCND3*, mirroring their expression patterns in normal neural development (Extended Data Fig. 7c). Furthermore, *CCND2* was upregulated in activated mouse NSCs before these cells enter the cell cycle (Extended Data Fig. 7d). These results suggest *CCND2* has a role in both normal and malignant stem cell programs.

Finally, we explored the role of genetic events in shaping cellular identity, devising two approaches to obtain genetic information from scRNA-seq and classify cells into tumour subclones. First, we used the inferred CNVs in each cell (Fig. 1a, b). Second, we defined subclonal point mutations from bulk DNA whole-exome sequencing, using ABSOLUTE<sup>25</sup>, and identified these mutations in the RNA-seq reads of individual cells, albeit with limited sensitivity (Methods).

In both analyses, we found genetic subclones that span all three of the compartments, although the genetic information obtained with these two approaches is partial and is not sufficient to reconstruct a full phylogenetic tree. We observed the same three sub-population architectures in distinct CNV subclones in MGH36 and in MGH97 (Fig. 1b), with cycling stem/progenitor cells and two lineages of differentiated non-cycling cells (Fig. 4a and Extended Data Fig. 8). Similarly, examining the distribution of expression states for cells harbouring subclonal point mutations, we found that 22 subclonal point mutations (Extended Data



**Figure 4 | Intra-tumoural genetic heterogeneity and association with gene expression states.** **a–d**, Cells were classified into genetic subclones based on CNVs (**a**, **b**) or *CIC* point mutation status (**c**, **d**), and examined for differences in gene expression states. **a**, Two CNV clones (green and grey) in MGH36 and MGH97 mapped to the cellular hierarchy defined by lineage (x axis) and stemness (y axis) scores. **b**, Percentages of cycling cells (x axis) and of stem/progenitor cells (y axis) in clone 1 (green) and clone 2 (grey) of MGH36 (square) and MGH97 (diamond). **c**, Cells were classified using mutation-specific qPCR as wild-type *CIC* (green), mutant *CIC* (orange) or *CIC* status not detected (black) and mapped to the cellular hierarchy. The fraction of mutant *CIC* cells as observed by qPCR (O) and as expected by ABSOLUTE (E) is indicated. **d**, An expression signature for mutant *CIC* cells. Shown is a heatmap of relative expression levels for *CIC*-dependent genes (rows) in mutant *CIC* cells (right) and wild-type *CIC* cells (left). Selected gene names are indicated.

Fig. 9) and a subclonal loss-of-heterozygosity event (Extended Data Fig. 10) are not significantly restricted to particular developmental states and often span all three states. Thus, the three compartments exist in different genetic subclones.

Although most subclonal mutations were of unknown functional relevance, we identified a subclonal mutation of *CIC* (~30% frequency in MGH53), a known tumour suppressor in oligodendroglioma<sup>26</sup>. RNA-seq reads detected the *CIC* mutation only in 7 MGH53 cells. We therefore designed a mutation-sensitive qPCR testing approach and were able to identify 28 mutant *CIC* cells (including all cells detected by RNA-seq) and 27 wild-type *CIC* cells (Fig. 4c). Notably, we identified a signature of expression changes between the mutant *CIC* and wild-type cells (Fig. 4d, Supplementary Table 5), including increased expression of *ETV1* and *ETV5* (ref. 27) in mutant *CIC* cells<sup>8</sup>. Despite these specific transcriptional changes, mutant *CIC* and wild-type *CIC* cells spanned all three subpopulations (Fig. 4c). Thus, many subclonal mutations can accrue within the hierarchy (but not drive it), although without a comprehensive phylogenetic reconstruction, we cannot categorically rule out a genetic influence.

Although CNV subclones in MGH36 and MGH97 included cells from all three tumour compartments, they differed in their relative

distributions (Fig. 4a, b and Extended Data Fig. 8). Clone 1 of MGH36 displayed a higher frequency of stem/progenitors, whereas clone 2 displayed higher frequency of astrocyte-like cells ( $P < 10^{-9}$ , Fisher's exact test). Similarly, clone 2 of MGH97 contained a higher frequency of stem/progenitors ( $P < 10^{-16}$ ). Furthermore, the frequencies of cycling cells were higher in clone 1 of MGH36 and clone 2 of MGH97, consistent with their increased frequencies of stem/progenitor cells. Thus, genetic evolution may modulate patterns of self-renewal and differentiation.

In conclusion, our results highlight the fact that there is a subpopulation of undifferentiated cells in oligodendroglioma that possess stem cell expression signatures and enriched proliferative potential. Thus, the most primitive and undifferentiated population of cancer cells might be the main source of proliferating cells in oligodendroglioma. Although we cannot rule out an influence from genetic mutations, many subclonal events span all three states, consistent with this architecture being primarily dictated by non-genetic developmental programs. A caveat of our work is that because grade II oligodendroglioma cells do not grow in xenotransplantation, we could not functionally validate the stem/progenitor program, and instead we infer its function from the inverse association with differentiation programs, the enriched proliferation and the similarity to normal stem/progenitors. Our single-cell profiles suggest that oligodendroglioma stem/progenitor cells more closely resemble a primitive tri-potent neural cell type, such as NSC or NPC than a more committed glial progenitor like an OPC<sup>17,28</sup>. By providing the genome-wide transcriptional signature of cancer stem/progenitor cells in oligodendroglioma, our work delineates cellular programs that represent promising targets to affect tumour growth. Further studies will be needed to functionally validate our findings, interrogate their generality across other glioma subtypes, and investigate opportunities for clinical translation.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 11 May; accepted 26 September 2016.**

**Published online 2 November 2016.**

- Kreso, A. & Dick, J. E. Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275–291 (2014).
- Lathia, J. D., Mack, S. C., Mulkearns-Hubert, E. E., Valentim, C. L. & Rich, J. N. Cancer stem cells in glioblastoma. *Genes Dev.* **29**, 1203–1217 (2015).
- Friedmann-Morvinski, D. et al. Dedifferentiation of neurons and astrocytes by oncogenes can induce gliomas in mice. *Science* **338**, 1080–1084 (2012).
- Louis, D. N., Ohgaki, H., Wiestler, O. D. & Cavenee, W. K. *WHO Classification of Tumors of the Central Nervous System* 4th edn (IARC, 2016).
- Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols* **9**, 171–181 (2014).
- Patel, A. P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).
- Butovsky, O. et al. Identification of a unique TGF- $\beta$ -dependent molecular and functional signature in microglia. *Nature Neurosci.* **17**, 131–143 (2014).
- Rousseau, A. et al. Expression of oligodendroglial and astrocytic lineage markers in diffuse gliomas: use of YKL-40, ApoE, ASCL1, and NKX2-2. *J. Neuropathol. Exp. Neurol.* **65**, 1149–1156 (2006).
- Zhang, Y. et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34**, 11929–11947 (2014).
- Feng, W. et al. The chromatin remodeler CHD7 regulates adult neurogenesis via activation of SoxC transcription factors. *Cell Stem Cell* **13**, 62–72 (2013).
- Ikushima, H. et al. Autocrine TGF- $\beta$  signaling maintains tumorigenicity of glioma-initiating cells through Sry-related HMG-box factors. *Cell Stem Cell* **5**, 504–514 (2009).
- Suvà, M. L. et al. Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. *Cell* **157**, 580–594 (2014).
- Rheinbay, E. et al. An aberrant transcription factor network essential for Wnt signaling and stem cell maintenance in glioblastoma. *Cell Reports* **3**, 1567–1579 (2013).
- Suvà, M. L., Riggi, N. & Bernstein, B. E. Epigenetic reprogramming in cancer. *Science* **339**, 1567–1570 (2013).
- Miller, J. A. et al. Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).
- Darmanis, S. et al. A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl Acad. Sci. USA* **112**, 7285–7290 (2015).
- Sugiarto, S. et al. Asymmetry-defective oligodendrocyte progenitors are glioma precursors. *Cancer Cell* **20**, 328–340 (2011).
- Zhang, Y. et al. Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. *Neuron* **89**, 37–53 (2016).
- Shin, J. et al. Single-cell RNA-seq with waterfall reveals molecular cascades underlying adult neurogenesis. *Cell Stem Cell* **17**, 360–372 (2015).
- Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
- Kowalczyk, M. S. et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res.* **25**, 1860–1872 (2015).
- The Cancer Genome Atlas Research Network Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N. Engl. J. Med.* **372**, 2481–2498 (2015).
- Lange, C. & Calegari, F. Cdks and cyclins link G1 length and differentiation of embryonic, neural and hematopoietic stem cells. *Cell Cycle* **9**, 1893–1900 (2010).
- Koyama-Nasu, R. et al. The critical role of cyclin D2 in cell cycle progression and tumorigenicity of glioblastoma stem cells. *Oncogene* **32**, 3840–3845 (2013).
- Carter, S. L. et al. Absolute quantification of somatic DNA alterations in human cancer. *Nature Biotechnol.* **30**, 413–421 (2012).
- Bettegowda, C. et al. Mutations in *CIC* and *FUBP1* contribute to human oligodendroglioma. *Science* **333**, 1453–1455 (2011).
- Padul, V., Epari, S., Moiyadi, A., Shetty, P. & Shirsat, N. V. ETV/Pea3 family transcription factor-encoding genes are overexpressed in *CIC*-mutant oligodendrogliomas. *Genes Chromosom. Cancer* **54**, 725–733 (2015).
- Liu, C. et al. Mosaic analysis with double markers reveals tumor cell of origin in glioma. *Cell* **146**, 209–221 (2011).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank L. Gaffney for graphic support. This work was supported by grants from the National Brain Tumor Society (to M.L.S. and D.N.L.), the Smith Family Foundation (to M.L.S.), NIH-NCI SP0RE on brain cancer Career Enhancement Project and Developmental Research Project (to M.L.S.), the Broad Institute Broadnext10 program (to M.L.S. and O.R.R.), the American Cancer Society (to M.L.S.) and start-up funds from the MGH department of Pathology. A.S.V. was supported by the NIH R25 fellowship (NS065743) and research grants from the American Brain Tumor Association and Neurosurgery Research and Education Foundation. I.T. was supported by a Human Frontier Science Program fellowship and a Rothschild fellowship. A.R. was supported by funds from the Howard Hughes Medicine Institute, the Klarman Cell Observatory, STARR cancer consortium, NCI grant 1U24CA180922, by the Koch Institute Support (core) grant P30-CA14051 from the National Cancer Institute, the Ludwig Center and the Broad Institute. A.R. is a scientific advisory board member for ThermoFisher Scientific and Syros Pharmaceuticals and a consultant for Driver Group. Flow cytometry and sorting services were supported by shared instrumentation grant 1S10RR023440-01A1. M.M. was supported by the California Institute of Regenerative Medicine (CIRM) grants RB4-06093 and RN3-06510 and the Virginia and D.K. Ludwig Fund for Cancer Research.

**Author Contributions** I.T., A.S.V., A.R. and M.L.S. conceived the project, designed the study, and interpreted results. A.S.V., C.H., L.E.E. and C.N. collected single cells and generated single-cell sequencing data. I.T. performed computational analyses. J.M.Fra, K.Y. and G.G. provided support for genomic and genetic analyses. J.M.Fis, C.R. and K.J.L. designed and performed qPCR experiments. C.C.L. and R.M. provided flow cytometry expertise. C.M. and M.M. developed normal human cell cultures used in the study. N.D., N.R., M.N.R., M.L.O. and A.J.I. performed *in situ* hybridization and FISH experiments. A.P.P., A.A.P., D.G., B.I., J.N., R.S., M.G.F., B.V.N., D.P.C., W.T.C., R.L.M., M.P.F., O.R.R., T.R.G., B.E.B. and D.N.L. provided experimental and analytical support. I.T., A.R. and M.L.S. wrote the manuscript with feedback from all authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.L.S. ([suva.mario@mgh.harvard.edu](mailto:suva.mario@mgh.harvard.edu)) or A.R. ([aregev@broadinstitute.org](mailto:aregev@broadinstitute.org)).

**Reviewer Information** Nature thanks P. Dirks, J. Rich and the other anonymous reviewer(s) for their contribution to the peer review of this work.



## METHODS

**Data reporting.** No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

**Tumour dissociation.** Patients at the Massachusetts General Hospital consented preoperatively to take part in the study in all cases following the Institutional Review Board Protocol 1999P008145. Fresh tumours were collected at time of resection and the presence of malignant cells was confirmed in frozen sections on adjacent, representative pieces of tissue. Fresh tumour tissue was minced with a scalpel and enzymatically dissociated using a gentle papain-based brain tumour dissociation kit (Miltenyi Biotec). Large pieces of debris were removed with a 100 µm strainer, and dissociated cells were layered carefully onto a 5 ml density gradient (Lympholyte-H, Cedar Lane labs), which was centrifuged at 2,000 r.p.m. for 10 min at room temperature to pellet dead cells and red blood cells. The interface containing live cells was saved and used for staining and flow cytometry. Viability was measured using trypan blue exclusion, which confirmed >90% cell viability.

**Fluorescence-activated cell sorting.** For primary tumour sorting, tumour cells were blocked in 1% bovine serum albumin in Hanks buffered saline solution (BSA / HBSS), and then stained first with CD45-Vioblue direct antibody conjugate (Miltenyi Biotec) for 30 min at 4 °C. Cells were washed with cold PBS, and then resuspended in 1 ml of BSA / HBSS containing 1 µM calcein AM (Life Technologies) and 0.33 µM TO-PRO-3 iodide (Life Technologies) to co-stain for 30 min before sorting. Fluorescence-activated cell sorting was performed on FACSaria Fusion Special Order System (Becton Dickinson) using 488 nm (calcein AM, 530/30 filter), 640 nm (TO-PRO-3, 670/14 filter), and 405 nm (Vioblue, 450/50 filter) lasers. Fluorescence-minus-one controls were included with all tumours, as well as heat-killed controls in early pilot experiments, which were crucial to ensure proper identification of the TO-PRO-3 positive compartment and ensure sorting of the live cell population. Standard, strict forward scatter height versus area criteria were used to discriminate doublets and gate only singlets. Viable cells were identified by staining positive with calcein AM but negative for TO-PRO-3. Single cells were sorted into 96-well plates containing cold buffer TCL buffer (Qiagen) containing 1% β-mercaptoethanol, snap frozen on dry ice, and then stored at -80 °C before whole transcriptome amplification, library preparation and sequencing.

**Whole-transcriptome amplification, library construction, sequencing, and processing.** Libraries from isolated single cells were generated based on the Smart-seq2 protocol (Picelli 2014) with the following modifications. RNA from single cells was first purified with Agencourt RNAClean XP beads (Beckman Coulter) before oligo-dT primed reverse transcription with Maxima reverse transcriptase and locked TSO oligonucleotide, which was followed by 20 cycle PCR amplification using KAPA HiFi HotStart ReadyMix (KAPA Biosystems) with subsequent Agencourt AMPure XP bead purification as described. Libraries were tagged using the Nextera XT Library Prep kit (Illumina) with custom barcode adapters (sequences available upon request). Libraries from 384 cells with unique barcodes were combined and sequenced using a NextSeq 500 sequencer (Illumina).

We also analysed 96 cells from MGH60 with an alternative protocol that incorporates random molecular tags (RMTs), also known as unique molecular identifiers, or UMIs) in order to control for PCR amplification bias, as described previously<sup>29</sup> and we obtained similar results.

Paired-end, 38-base reads were mapped to the UCSC hg19 human transcriptome using Bowtie with parameters “-q-phred33-quals -n 1 -e 99999999 -l 25 -l 1 -X 2000 -a -m 15 -S -p 6”, which allows alignment of sequences with single base changes, such as point mutation in the *IDH1* gene. Expression values were calculated from SAM files using RSEM v1.2.3 in paired-end mode using parameters “-estimate-rspd-paired-end -sam -p 6”, from which TPM values for each gene were extracted.

**Immunohistochemistry.** Haematoxylin and eosin and single antibody staining (GFAP, Ki-67) was done by the clinical pathology laboratory at the Massachusetts General Hospital per routine protocol. For GFAP and Ki-67 double immunohistochemistry, paraffin-embedded sections were mounted on glass slides, deparaffinized in xylene, treated with 0.5% peroxide in methanol, and rehydrated. Antigen retrieval was done using sodium citrate-based, heat-induced antigen retrieval at pH 6.0. The Dako EnVision G/2 double stain system was used for blocking, staining, and development using rabbit anti-Ki67 antibody (Abcam ab15580 at 1:300) and mouse anti-GFAP antibody (Dako M0761 at 1:100).

**RNA *in situ* hybridization.** Human tissue was obtained from the Massachusetts General Hospital according to an Institutional Review Board-approved protocol (1999P008145) and informed consent was obtained from all patients. ViewRNA technology (Affymetrix) was used for manual format RNA *in situ* hybridization. Tissue sections mounted on glass slides were stored at -80 °C until they were used for hybridization. Slides were baked at 60 °C for 1 h, then denatured at 80 °C for 3 min, deparaffinized with HistoClear and ethanol dehydration. RNA targets in

dewaxed sections were unmasked by treating with pre-treatment buffer at 95 °C for 10 min and digested with 1:100 dilution protease at 40 °C for 10 min, followed by fixation with 10% formalin for 5 min at room temperature. Probe concentrations were 1:40 for both type 1 (red) and type 6 (blue) probe sets, except that the ApoE probe was used at 1:80 dilution. The probe was incubated on sections for 2 h at 40 °C and then washed serially. Affymetrix Panomics probes included ApoE (type 6, catalogue number VA6-16904 and type 1, catalogue number VA1-18265), OMG (type 1, catalogue number VA1-18161), Sox4 (type 6, catalogue number VA6-18162), CCND2 (type 6, catalogue number VA6-18266), Ki-67 (type 1, catalogue number VA1-11033). Signal was amplified using PreAmplifier mix QT for 25 min at 40 °C followed by Amplifier mix QT for 15 min at 40 °C, and then signal was hybridized with labelled probe at 1:1,000 dilution for 15 min at 40 °C. Colour was developed using Fast Blue substrate for Type 6 probes and Fast Red substrate for Type 1 probes for 30 min at 40 °C. Tissue was counterstained with Gill's haematoxylin for 25 s at room temperature followed by mounting with ADVANTAGE mounting media (Innovex). For quantification of compartments by ISH, at least 1,000 cells were counted in representative areas of the tumours.

**Fluorescent *in situ* hybridization (FISH).** The probes used in this study consisted of centromeric (CEP) and locus-specific identifiers (LSI) probes. CEP probes included: CEP2 (2p11.1-q11.1, spectrum orange), CEP4 (4p11-q11, spectrum aqua), CEP9 (9p11-q11, spectrum aqua), CEP12 (12p11.1-q11, spectrum green), CEP17 (17p11.1-q11.1, spectrum aqua) and Y (Yp11.1-q11.1, spectrum green) all obtained from Abbott Molecular, Inc. (Des Plaines, IL). LSI probes were 1p36/1q25 and 19q13/19p13 dual-colour probe set (Abbott), and bacterial artificial chromosome RP11-351D16 (10q11.21, spectrum red or green; CHORI, Oakland, CA).

FISH was performed as described previously<sup>30</sup>. Briefly, 5-µm sections of formalin-fixed, paraffin-embedded tumour material were deparaffinized, hydrated, and pretreated with 0.1% pepsin for 1 h. Slides were then washed in 2× saline-sodium citrate buffer (SSC), dehydrated, air dried, and co-denatured at 80 °C for 5 min with a three-colour probe panel and hybridized at 37 °C overnight using the Hybrite Hybridization System (Abbott). Two 2-min post-hybridization washes were performed in 2× SSC/0.3%NP40 at 72 °C followed by one 1-min wash in 2× SSC at room temperature. Slides were mounted with Vectashield containing 4',6-diamidino-2-phenylindole (Vector, Burlingame, CA, USA). Entire sections were observed with an Olympus BX61 fluorescent microscope equipped with a charge-coupled device camera and analysed with Cytovision software (Applied Imaging, Santa Clara, CA).

**Human NPC culturing.** Human NPCs were dissociated from the subventricular zone of 19 week fetal tissue and resulting neurospheres were expanded in a 1:1 mixture of DMEM/F12 and Neurobasal A (Invitrogen), supplemented with B27 lacking vitamin A, EGF, FGF, and heparin. Single live NPCs were isolated by FACS from a passage 8 culture and sorted into 96-well plates containing Buffer TCL (Qiagen) + 1% β-mercaptoethanol. For differentiation assays, NPCs were plated in chamber slides coated with poly-D-lysine and laminin, and proliferation media was exchanged over a period of 3 days with base media supplemented with either 1% FBS, 1% FBS + 60 ng ml<sup>-1</sup> T3, or FBS + 100 nM trans-retinoic acid and 10 ng ml<sup>-1</sup> NT3. Multipotency was confirmed by indirect immunofluorescence after 7 days of differentiation with GFAP (Abcam ab53554), Olig2 (Millipore AB9610), and Neurofilament (Aves).

**Single-cell RNA-seq data processing.** Expression levels were quantified as  $E_{ij} = \log_2(TPM_{ij}/10 + 1)$ , where  $TPM_{ij}$  refers to transcript-per-million for gene  $i$  in sample  $j$ , as calculated by RSEM<sup>31</sup>. TPM values are divided by 10, since we estimate the complexity of single-cell libraries in the order of 100,000 transcripts and would like to avoid counting each transcript ~10 times, as would be the case with TPM, which may inflate the difference between the expression level of a gene in cells in which the gene is detected and those in which it is not detected.

For each cell, we quantified two quality measures: the number of genes for which at least one read was mapped, and the average expression level of a curated list of housekeeping genes. We then conservatively excluded all cells with either fewer than 3,000 detected genes or an average housekeeping expression ( $E$ , as defined above) below 2.5. For the remaining cells we calculated the aggregate expression of each gene as  $\log_2(\text{average}(TPM_{i,1..n}) + 1)$ , and excluded genes with an aggregate expression below 4, leaving a set of 8,008 analysed genes. For the remaining cells and genes, we defined relative expression by centering the expression levels,  $Er_{ij} = E_{ij} - \text{average}[E_{i,1..n}]$ . Centring was performed within each tumour separately in order to decrease the impact of inter-tumoural variability on the combined analysis across tumours.

**CNV estimation.** Initial CNVs ( $CNV_0$ ) were estimated by sorting the analysed genes by their chromosomal location and applying a moving average to the relative expression values, with a sliding window of 100 genes within each chromosome, as previously described<sup>6</sup>. To avoid considerable impact of any particular gene on the moving average, we limited the relative expression values to  $[-3, 3]$  by replacing



all values above 3 by 3, and replacing values below  $-3$  by  $-3$ . This was performed only in the context of CNV estimation. For visualization purposes, in order to include the two chromosomes with fewest analysed genes (chromosome 18 and 21 with 105 and 75 genes, respectively), we extended the moving average to include up to 50 genes from the flanking chromosomes (for example, the first window in chromosome 18 consisted of the last 50 genes of chromosome 17 and the first 50 genes of chromosome 18, whereas the 51 through 56 windows in that chromosome consisted only of chromosome 18 genes). This initial analysis is based on the average expression of genes in each cell compared to the other cells and therefore does not have a proper reference which is required to define the baseline. However, we detected a cluster of cells that have higher values at chromosome 1p and 19q, which we know are deleted in the six tumours, and that have consistent 'CNV patterns' across the genome, despite the fact that they originate from all three tumours. We thus defined these as the non-cancer cells and used the average CNV estimate at each gene across these cells as the baseline. The non-cancer cells included both microglia and oligodendrocytes, which differed in gene expression patterns and therefore also in CNV estimates (for example, the MHC region in chromosome 6 had consistently higher values in microglia than in oligodendrocytes and cancer cells). We therefore defined two baselines, as the average of all microglia and the average of all oligodendrocytes, and based on these the maximal (*BaseMax*) and minimal (*BaseMin*) baseline at each genomic position. The final CNV estimate of cell  $i$  at position  $j$  was defined as:

$$CNV_f(i, j) = \begin{cases} CNV_0(i, j) - BaseMax(j), & \text{if } CNV_0(i, j) > BaseMax(j) + 0.2 \\ CNV_0(i, j) - BaseMin(j), & \text{if } CNV_0(i, j) < BaseMin(j) - 0.2 \\ 0, & \text{if } BaseMin(j) - 0.2 < CNV_0(i, j) < BaseMin(j) + 0.2 \end{cases}$$

**Principal component analysis.** We performed principal component analysis (PCA) for the relative expression values of all cancer cells (as defined by CNV analysis) from the three tumours combined. The covariance matrix used for PCA was generated using an approach outlined in ref. 32 to decrease the weight of less reliable 'missing' values in the data. The basis of this approach is that due to the limited sensitivity of single-cell RNA-seq, many genes are not detected in particular cells despite being expressed. This is particularly pronounced for genes expressed at low levels, and for cells with low library complexity (that is, for which relatively few genes are detected), and results in non-random patterns in the data, whereby cells may cluster based on their complexity and genes may cluster based on their expression levels, rather than 'true' co-variation. To mitigate this effect, we assign weights to missing values, such that the weight of  $E_{ij}$  is proportional to the expectation that gene  $i$  will be detected in cell  $j$  given the average expression of gene  $i$  and the total complexity (number of detected genes) of cell  $j$ .

To further verify that the PCA results are not driven by library complexity, we compared the PCA results to those of shuffled data. We iteratively swapped the expression of individual genes between pairs of cells with similar complexities, swapping each gene in each cell at least once. In that way we shuffled the data and removed the biological clustering, but maintained the distribution of complexities across cells, as well as the distribution of expression levels for each gene. PCA over the shuffled data defined the complexity-based effect, as evident by a Pearson correlation of 0.96 between the PC1 cell scores and their complexities (in the original data this correlation is only 0.41). We then compared PC1 gene scores between the original and the shuffled data (Extended Data Fig. 2f). Although PC1 gene scores of most genes are comparable between the two analyses, the loadings of the oligodendrocyte and astrocyte gene sets (Supplementary Table 1) were highly affected. Oligodendrocyte genes were originally associated with highly positive PC1 scores, and their scores are significantly decreased upon shuffling (97% of the oligodendroglial genes were among the 5% genes with the most decreased loadings,  $P < 10^{-32}$ ); similarly, astrocytic genes were originally associated with negative PC1 scores, and their scores are significantly increased upon shuffling (all astrocytic genes were among the 5% of genes with the most increased loadings,  $P < 10^{-32}$ ). As a result, none of the genes with highest and lowest PC1 scores (after shuffling) overlap with our oligodendroglial and astrocytic gene sets. Thus, complexity does not account for the association of PC1 with the differentiation programs. Similarly, complexity clearly does not account for the PC2 and PC3 stemness program, as PC2 cell scores are positively correlated with complexity ( $R = 0.27$ ), whereas PC3 cell scores are negatively correlated with complexity ( $R = -0.24$ ) and stemness genes were defined as those associated with both PC2 and PC3.

**PC1-associated genes and lineage scores.** The top correlated genes with PC1 scores (across all tumour cells) were defined as PC1-associated genes. We focused on the genes with an absolute correlation value above 0.35, but note that other thresholds gave similar results (not shown). Of those genes, the subset that was differentially expressed by at least threefold between oligodendrocyte (OC) and astrocyte (AC) mouse cells<sup>9</sup>, and for which the two comparisons were consistent

(that is, PC1-positively correlated genes with higher OC expression, and PC1-negatively correlated genes with higher AC expression) were defined as the OC and AC lineage gene sets. Lineage scores were then calculated as the average relative expression of the lineage gene set minus the average relative expression of a control gene set, that is,  $Lin_{ij} = \text{average}[Er(G_{ij})] - \text{average}[Er(G_{ij}^{cont})]$ , in which  $Lin_{ij}$  is the score of cell  $i$  to lineage  $j$ ,  $G_{ij}$  is the gene set for lineage  $j$  and  $G_{ij}^{cont}$  is a control gene set for lineage  $j$ . The control gene set was defined by first binning all 8,008 analysed genes into 25 bins of aggregate expression levels and then, for each gene in the lineage gene set, randomly selecting 100 genes from the same expression bin. In this way, the control gene set has a comparable distribution of expression levels to that of the lineage gene set and the control gene set is 100-fold larger, such that its average expression is analogous to averaging over 100 randomly selected gene sets of the same size as the lineage gene set. The final lineage score of each cell was defined as the maximal score over the two lineages,  $LIN_i = \max(Lin_i, OC, AC)$ . For visualization purposes where the two lineage scores are shown in a single axis, we first assigned random scores within (0–0.15) to all cells with  $LIN < 0$ , to avoid having many overlapping cells at  $x = 0$ . Second, we assigned negative scores to the cells with higher AC than OC scores (that is, a cell with AC and OC scores of 0.1 and 1, respectively, would be assigned a lineage score of 1, whereas a cell with AC and OC scores of 1 and 0.1 would be assigned a lineage score of  $-1$ ).

**PC2 and PC3 associated genes and stemness scores.** Both PC2 and PC3 were associated with intermediate values of PC1 (Extended Data Fig. 2c) and therefore with presumably less differentiated cells, and both were correlated with a shared set of genes, but distinguished by their correlation with cell 'complexity'. We considered their sum as a potential stemness program. To detect potential stem-related genes, we chose the top 100 most positively correlated genes with PC2 + PC3 scores across all cancer cells from the three tumours. The 100 candidate genes were then restricted to (1) genes that are positively correlated with both PC2 and PC3, which primarily excluded ribosomal protein genes that were only correlated with PC2; (2) genes for which the average relative expression among the stem-like cells was above average. Stemness scores for each cell,  $stem(i)$ , were then defined as the average relative expression of the stemness gene-set ( $G_{stem}$ ) minus the average of a control gene set ( $G_{stem}^{cont}$ ) and minus the lineage score of cell  $i$ :

$$Stem(i) = \text{average}[Er(G_{stem})] - \text{average}[Er(G_{stem}^{cont})] - LIN(i)$$

**Assignment of cells to four subpopulations: stem/progenitor-like, undifferentiated, OC-like and AC-like.** Cells were scored for the three programs defined above (two lineage scores and a stemness score) and assigned to the subpopulation that corresponds to their highest scoring program, if the maximal score was above 0.5 and was higher by 0.5 than the score for the other programs. Cells in which the maximal score did not pass these thresholds were assigned to the undifferentiated subpopulation, for which we did not detect a specific expression program. We note that the expression programs are continuous and thus it is difficult to assign every cell to a discrete subpopulation. Nevertheless, most cells are highly biased towards one of the three states, and the overall estimates are consistent between analysis of single-cell RNA-seq data and tissue staining experiments (Extended Data Fig. 8c and Supplementary Table 3). Furthermore, very few cells ( $\sim 1\%$  on average, and 5% at most) scored for two programs simultaneously (with the same threshold of 0.5, Supplementary Table 3).

**Cell cycle analysis.** Analysis of single-cell RNA-seq in human (293T) and mouse (3T3) cell lines<sup>20</sup>, and in mouse haematopoietic stem cells<sup>21</sup> revealed in each case two prominent cell cycle expression programs that overlap considerably with genes that are known to function in replication and mitosis, respectively, and that have also been found to be expressed at G1/S phases and G2/M phases, respectively, in bulk samples of synchronized HeLa cells<sup>33</sup>. We thus defined a core set of 43 G1/S and 55 G2/M genes that included those genes that were detected in the corresponding expression clusters in all four datasets from the three studies described above (Supplementary Table 2). As expected, the genes in each of those expression programs were highly co-regulated in a small fraction of the oligodendroglial cells, such that some cells expressed only the G1/S or the G2/M programs and other cells expressed both programs (Extended Data Fig. 6a). Plotting the average expression of these programs revealed an approximate circle (Fig. 3a), which we hypothesize describes the progression along the cell cycle. Putative cycling cells were identified by at least a twofold upregulation and a  $t$ -test  $P$  value  $< 0.01$  for either the G1/S or the G2/M gene set compared to the average of all cells. Although we cannot confidently define the regions that correspond to each phase of the cell cycle in an automatic way, we manually defined four regions in the apparent circle and assigned them to approximate cell cycle phases.

**Analysis of whole-exome DNA sequencing data.** Output from Illumina software was processed by the Picard processing pipeline to yield BAM files containing aligned reads (bwa version 0.5.9, to the NCBI Human Reference Genome Build

hg19) with well-calibrated quality scores<sup>34,35</sup>. Sample contamination by DNA originating from a different individual was assessed using ContEst<sup>36</sup>. Somatic single nucleotide variations (sSNVs) were then detected using MuTect<sup>37</sup>. Following this standard procedure, we filter sSNVs by (1) removing potential DNA oxidation artefacts<sup>38</sup>; (2) removing events seen in sequencing data of a large panel of ~8,000 TCGA normal samples; (3) realigning identified sSNVs with NovoAlign (<http://www.novocraft.com>) and performing an additional iteration of MuTect with the newly aligned BAM files. sSNVs were finally annotated using Oncotator<sup>39</sup>. Sample purity and ploidy, as well as Cancer Cell Fraction (CCF) of identified sSNVs were determined by ABSOLUTE<sup>25</sup>. Genome-wide copy-ratio profiles were inferred using CapSeg. Read depth at capture targets in tumour samples was calibrated to estimate copy ratio using the depths observed in a panel of normal genomes. Next, we performed allelic copy analysis using reference and alternate counts at germline heterozygous SNP sites.

**Mutation calling in single cells.** sSNVs that were identified by WES were examined in single-cell RNA-seq data by the mpileup command of SAMtools<sup>40</sup>. The fraction of cells in which we identified these mutations was, on average, only 1.3% of the expected fraction estimated by ABSOLUTE. This low sensitivity primarily reflects the low coverage of the RNA-seq reads over the transcriptome of single cells. Accordingly, sensitivity was correlated with the expression levels of the genes that harbour the mutations, and reached 20.4% for the top 10% most highly expressed genes. Sensitivity was also affected by heterozygosity and allele-specific expression, as in some heterozygote mutant cells we might only sequence the wild-type allele.

We used a targeted sequencing approach to increase our sensitivity for three specific mutations in MGH54 which were identified by WES but detected in very few cells by single-cell RNA-seq. We designed primers flanking these three mutations (in ZEB2, EE1B2 and DNJC4), PCR-amplified single cell cDNAs (frozen stocks of product from the pre-amplification reaction of the Smart-seq2 protocol) and sequenced the amplified material. This approach was applied for 1,056 cells from MGH54. Mutant cells were defined as those with at least 50 reads that mapped to the mutant allele as defined by WES, and for which the fraction of mutant reads was at least 20% of all reads and fivefold higher than the overall rate of mutant reads (in order to exclude a low rate of mutant reads due to PCR or sequencing errors). The mutations detected by this criteria were highly consistent with those identified from single-cell RNA-seq ( $P < 10^{-5}$ , hypergeometric test) and uncovered 19 additional mutant calls (three for ZEB2, three for EE1B2 and 13 for DNJC4).

We next focused on the 23 subclonal mutations for which (1) the estimated clonal fraction by ABSOLUTE was at most 60%; (2) at least three cells were identified as harbouring the mutation; and (3) at least one cell was identified as having a wild-type allele of the mutant gene. For each of those 23 mutations we plotted the lineage and stemness scores of all mutant cells to examine their distribution of expression states (Fig. 4 and Extended Data Fig. 9). Note that for these mutations we detected on average 9.4% of the expected fraction by ABSOLUTE.

To estimate the frequency of false-positive errors we defined, for each mutation that is detected by WES and analysed by RNA-seq mutation calling, (i) 'expected mutations': the number of events in which we find the exact mutation reported by WES; and (ii) 'false mutations', the number of events in which we find a mismatch in the same exact site but to a different base than expected by WES (there are 2 such possible bases). This approach focuses on the exact genomic context of the real mutations to obtain a reliable estimate of the false positive rate. This estimate is half the number of false mutations divided by the number of expected mutations (given 4 bases, one of which is wild type, there are two types of false mutations but only one type of expected mutations). The result of this analysis was an estimated average false positive rate of 0.85%, suggesting that the confidence of each detected mutation is, on average, higher than 99%. Accordingly, even in the most extreme case (for example, ZEB2) where only a single mutant cell is detected in one of the compartments of the hierarchy, we still have a 99% confidence that the mutation is represented in that compartment.

**Mutation-detecting qPCR and analysis of *CIC* mutations.** To detect *CIC* mutations in single cells from MGH53, we performed qPCR using SuperSelective PCR primers, which are highly specific to single base changes due to a loop-out sequence adjacent to the mutant base (<http://legacy.labroots.com/user/webinars/details/id/95>). The following qPCR primers were designed to target the c.4543 C > T, p.1515 R > C mutation on *CIC* cDNA which had been identified as subclonal in MGH53 via whole-exome sequencing analysis. Wild-type-specific forward primer: 5'-CCCTCCAAGGTTTGTCTGCAGccattcGAGGTGC-3'; mutant-specific forward primer: 5'-CCCTCCAAGGTTTGTCTGCAGccattcGAGGTGT-3'; universal reverse primer: 5'-tcgGGCAGCCTGCATGATCTT-3'.

The specificity of the single cell qPCR primers was validated by two approaches: (1) qPCR on artificial templates differing by only the mutant base; and (2) qPCR on

cDNA of single MGH53 tumour cells for which RNA-seq already detected mutant or wild-type reads. These positive control reactions were highly consistent between duplicates and with the mutation status as inferred from RNA-seq: qPCR identified 7 out of 7 mutant cells and 12 out of 15 wild-type cells, while the remaining three cells had no qPCR signal, and therefore all qPCR signal was consistent with RNA-seq data. We also took advantage of the fact that *CIC* is located on chr19q which is deleted in MGH53 cancer cells and therefore each cell only contains one *CIC* allele (loss-of-heterozygosity, LOH). Thus, in a single MGH53 cancer cell, we expect evidence of either mutant or wild-type *CIC*, but not both. Indeed, all cells with a signal in the positive control assay showed a difference in  $C_t$  values of at least 5 between mutant and wild-type reactions, consistent with LOH.

cDNA was taken from frozen stocks of product from the preamplification reaction of the Smartseq2 protocol. 1 µl from each well of cDNA was used as template for a second round of Smartseq2 preamplification and bead purification in order to increase overall signal downstream. qPCR was performed with the Fast Plus EvaGreen qPCR Master Mix Low Rox (Biotium 31014-1) according to the manufacturer's instructions with the sole modification of adding EDTA to a final reaction concentration of 1.6 mM to enhance primer selectivity.  $C_p \geq 33$  were considered negative signal;  $C_p < 33$  was considered positive signal.

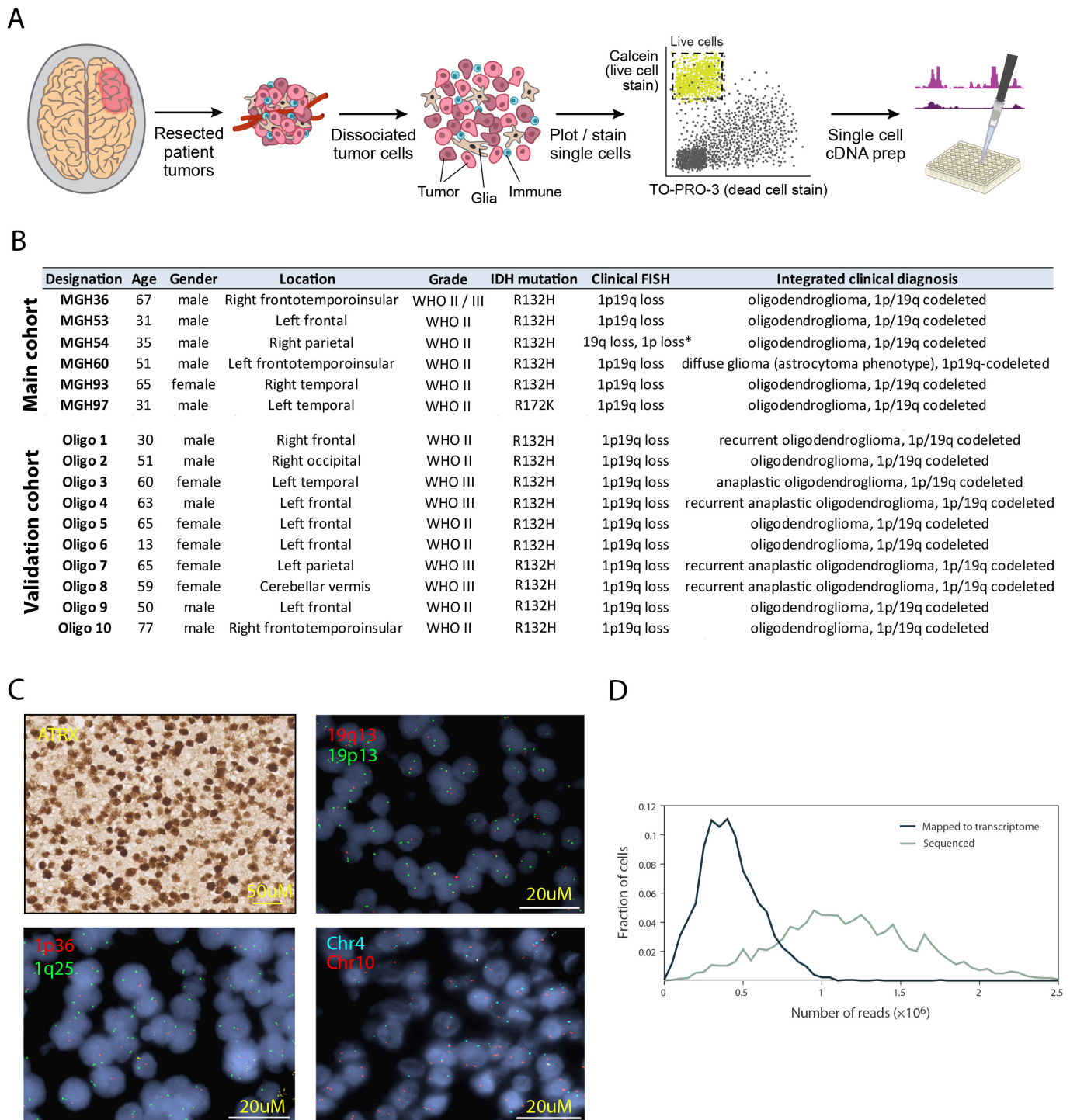
We performed SuperSelective qPCR on cDNA from 467 single MGH53 tumour cells. Of these, 61 cells had signal in both replicates for either mutant or wild type primers, but never for both. These were used to define 28 mutant *CIC* cells and 27 wild-type *CIC* cells, after excluding 6 cells which did not pass the single cell RNA-seq quality control filters.

To identify genes regulated by the *CIC* mutation, we compared the 28 mutant *CIC* cells and 27 wild-type *CIC* cells and identified genes with at least twofold average expression difference and  $P < 0.01$  (before correction for multiple hypothesis testing) based both on a permutation test and a  $t$ -test. To further filter the list of differentially expressed genes we also compared the mutant *CIC* cells to the 671 unresolved cells (in which we did not detect signal for either mutant or wild-type alleles by qPCR and by RNA-seq). As the fraction of *CIC* mutants was estimated as 30% by ABSOLUTE, we expect the unresolved cells to be a mixture of about one-third mutant *CIC* cells and about two-thirds wild-type *CIC* cells, and thus *CIC*-regulated genes should also differ between this mixture and the *CIC* mutants but to a lesser extent; we used a threshold of 1.5-fold difference between the average expression in *CIC* mutants and in unresolved cells. The resulting set of differentially expressed genes is given in Supplementary Table 5. We simulated this analysis with 1,000 randomly selected sets of cells (to replace the mutant *CIC* and wild-type *CIC* cells) and found an average of only five upregulated genes by the same criteria, suggesting a false discovery rate lower than 0.1 for the genes upregulated by the *CIC* mutation.

**Data availability.** Data generated for this study are available through the Gene Expression Omnibus (GEO) under accession number GSE70630.

29. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnol.* **33**, 495–502 (2015).
30. Mohapatra, G. *et al.* Glioma test array for use with formalin-fixed, paraffin-embedded tissue: array comparative genomic hybridization correlates with loss of heterozygosity and fluorescence *in situ* hybridization. *J. Mol. Diagn.* **8**, 268–276 (2006).
31. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
32. Shalek, A. K. *et al.* Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 363–369 (2014).
33. Whitfield, M. L. *et al.* Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* **13**, 1977–2000 (2002).
34. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
35. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
36. Cibulskis, K. *et al.* ContEst: estimating cross-contamination of human samples in next-generation sequencing data. *Bioinformatics* **27**, 2601–2602 (2011).
37. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnol.* **31**, 213–219 (2013).
38. Costello, M. *et al.* Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Res.* **41**, e67 (2013).
39. Ramos, A. H. *et al.* Oncotator: cancer variant annotation tool. *Hum. Mutat.* **36**, E2423–E2429 (2015).
40. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

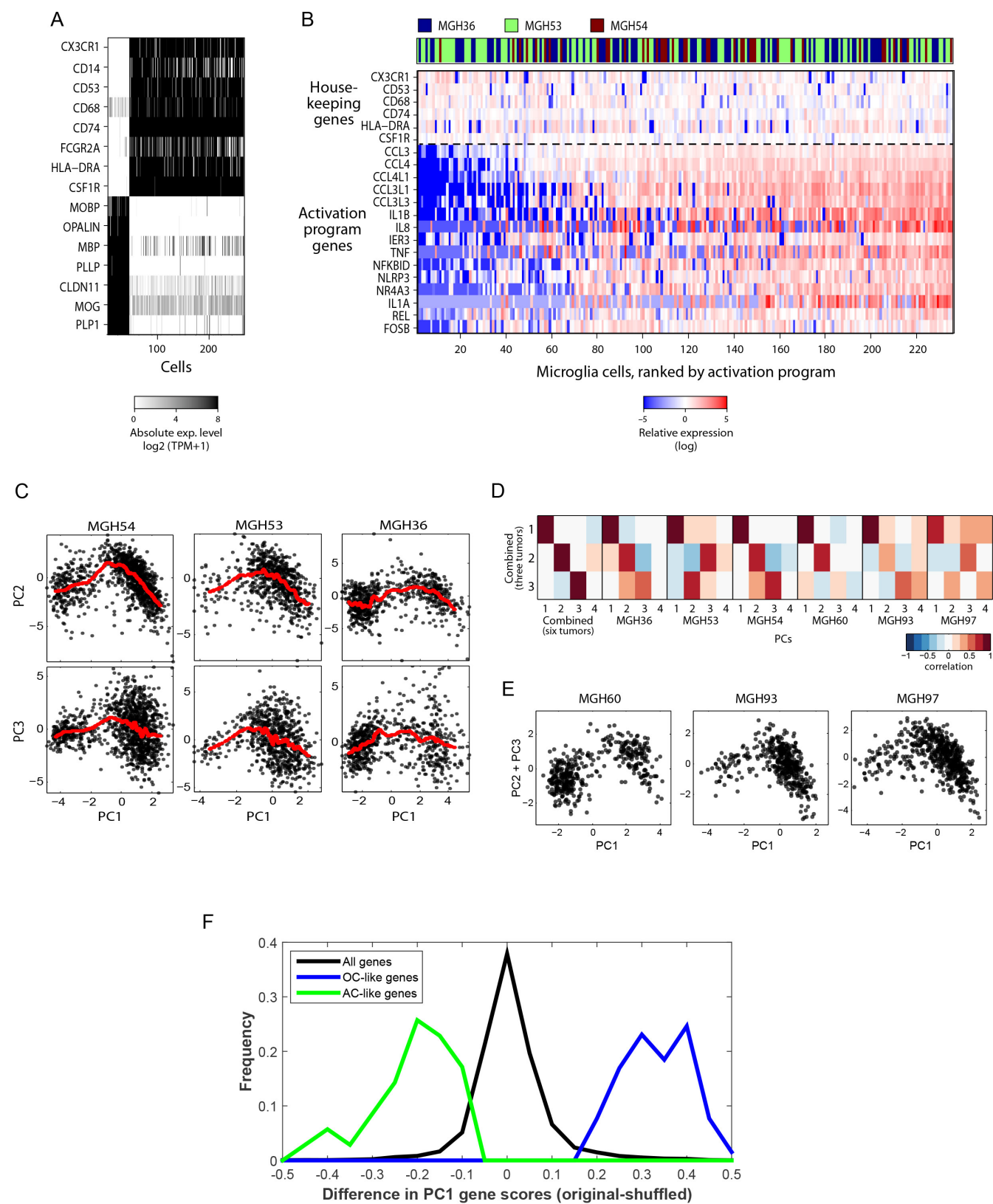




**Extended Data Figure 1 | Single-cell RNA-seq analysis of human oligodendroglioma samples.** **a**, Experimental workflow. **b**, Clinical information of the main and validation patient cohorts analysed in this study. Asterisk indicates a borderline result of chromosome 1p loss based on clinical testing. **c**, ISH (top left) and FISH (all other panels) in a representative tumour (MGH36). All our cases retain ATRX protein expression by ISH (top left) and show loss of chromosomes arms 1p

(bottom left) and 19q (top right) by FISH. In addition, tumour-specific CNVs identified by single-cell RNA-seq were confirmed by FISH (for example, loss of chromosome 4 in MGH36, bottom right panel). **d**, Distributions of the total number of sequenced paired-end reads per cell (grey) and of paired-end reads that were mapped to the transcriptome and used to quantify gene expression (black).



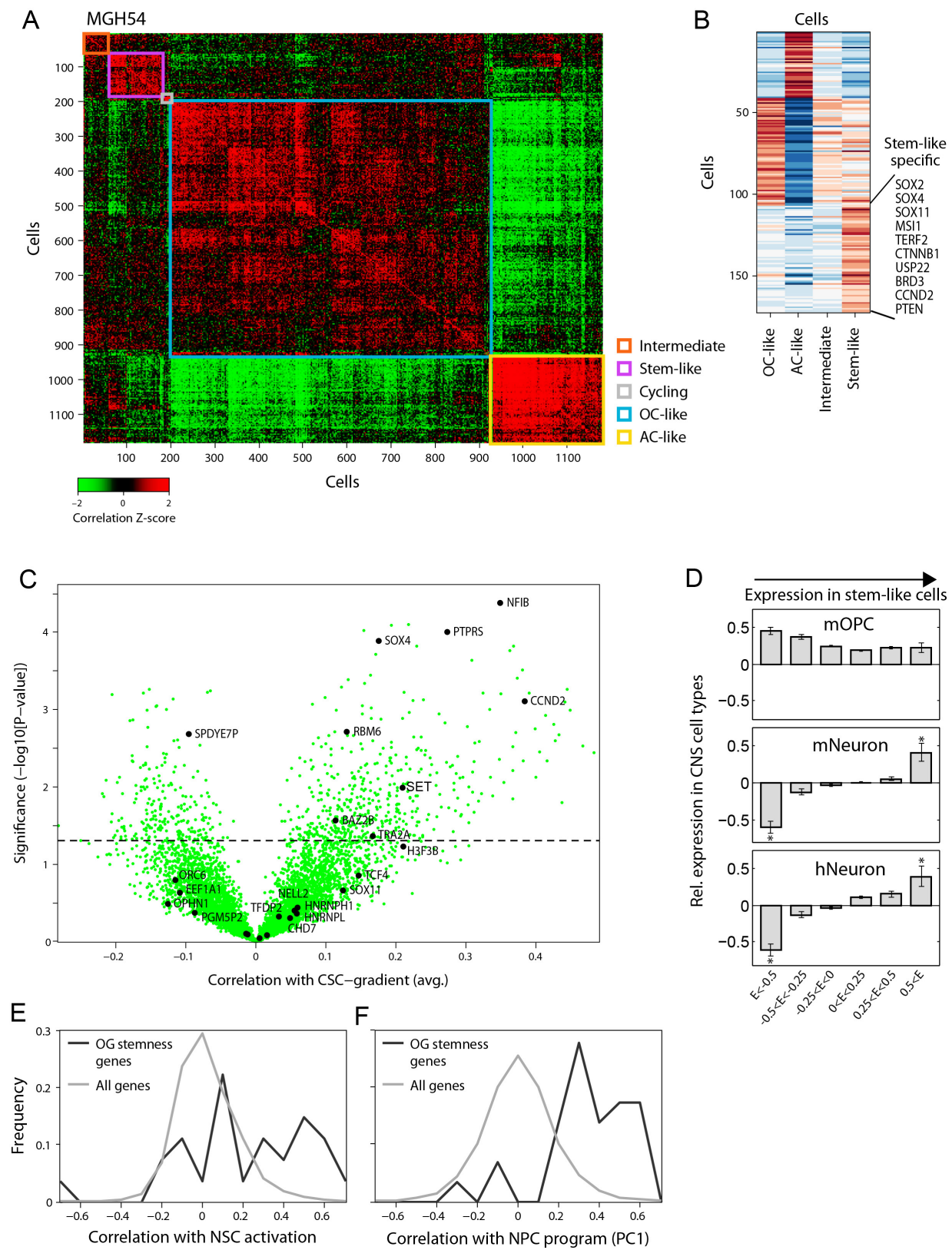


Extended Data Figure 2 | See next page for caption.

**Extended Data Figure 2 | Diversity of expression programs in oligodendrogloma.**

**a**, Two populations of non-cancer cells identified in oligodendrogloma. Selected genes that are differentially expressed among the two populations of non-cancer cells that lack CNVs (Fig. 1b, top), including markers of microglia (top) and oligodendrocytes (bottom). **b**, Expression programs in microglia cells from three tumours. The heat map shows relative expression of genes (rows) across microglia cells (columns). Above the dashed line are microglia markers expressed in all microglia cells and below the line are the genes of a microglia activation program, which is variably expressed, and includes cytokines, chemokines, early response genes and other immune effectors. This latter gene set might reflect a microglia activation program that could either be a general microglia program or potentially specific to the context of oligodendrogloma. Microglia cells ( $n = 235$ ) (columns) are rank ordered by their relative expression of the activation program. The tumour of origin of each cell is colour-coded as indicated in the top row. **c**, PC2 and PC3 are associated with intermediate values of PC1. PC1 scores are shown along with PC2 (top) and PC3 (bottom) scores for cells in each of the three tumours profiled at high depth. The red line indicates local weighted regression (LOWESS) with a span of 5%, which demonstrates

that PC2 and PC3 values tend to be highest in intermediate values of PC1 and to decrease in either high PC1 (that is, oligodendrocyte-like cells) or low PC1 (that is, astrocyte-like cells). **d**, Consistency of PCA across tumours. Shown are the Pearson correlations in gene loadings (over all analysed genes) between the top three PCs in PCA of the three tumours profiled at high depth ( $y$  axis, as shown in Fig. 1) and the top four PCs in alternative PCA of either all six tumours (left), as well as of PCA of each individual tumour (right). PC1–3 are highly consistent between the three-tumour and six-tumour PCAs ( $R > 0.9$ ); PC1 is highly consistent ( $R > 0.8$ ) between the three-tumour analysis and all other analysis. **e**, PC1 ( $x$  axis) and PC2 plus PC3 ( $y$  axis) scores of malignant cells from each of the three tumours profiled at intermediate depth, showing consistent patterns with those shown in Fig. 1d. **f**, Distribution of differences in PC1 loadings between the original PCA and the shuffled PCA (see description in the Methods section for principal component analysis) for all genes (black), oligodendrocyte-like (OC-like) genes (blue) and astrocyte-like (AC-like) genes (green). This analysis demonstrates that oligodendrocyte-like and astrocyte-like gene sets are highly skewed in the original PCA and their loadings are not recapitulated by shuffled data reflecting the effect of complexity.



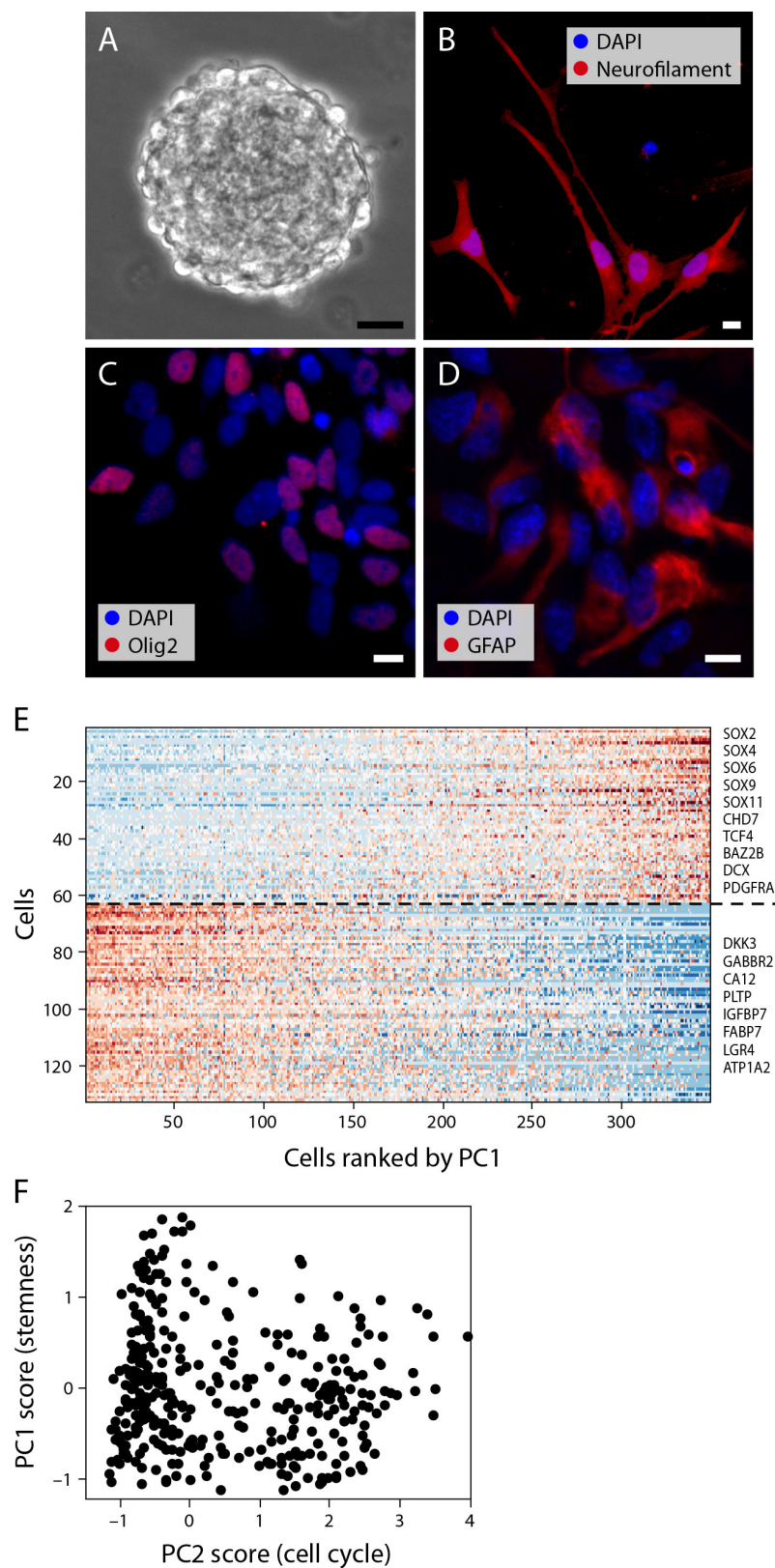
Extended Data Figure 3 | See next page for caption.



**Extended Data Figure 3 | The stemness program in oligodendrogloma.**

**a**, Cell–cell correlation matrix based on all analysed genes across all malignant cells in MGH54 ( $n = 1,174$ ). Cells are ordered by average linkage hierarchical clustering, and coloured boxes indicate distinct clusters. Clusters are marked based on the identity of differentially expressed genes as OC-like (blue), AC-like (yellow), cycling (pink) stem-like (purple) and intermediate cells that do not score highly for any of those expression programs (orange). **b**, Most differentially expressed genes. Shown is the average expression in each of the OC-like, AC-like, stem-like and intermediate cell clusters (columns) of differentially expressed genes (rows) defined by comparing cells from each of the OC-like, AC-like and stem-like clusters to cells from the remaining clusters with a  $t$ -test. Similar genes are highlighted as in Fig. 1 (OC-like: *OMG*, *OLIG1*, *OLIG2*, *SOX8*; AC-like: *ALDOC*, *APOE*, *SOX9*; Stem-like: *SOX4*, *SOX11*, *CCND2*, *SOX2*). Stem-like genes also include *CTNNB1*, *USP22*, and *MSI1*. **c**, Overlap with human GBM stemness program. We previously<sup>6</sup> identified a GBM stemness program and determined the association of each gene with that program by the correlation between the expression of that gene and the average expression of the stemness program's genes across individual cells ('CSC gradient') in each of five GBM tumours. Shown is the average correlation ( $x$  axis) of each analysed gene (green dots) across the five cases and the  $P$  values of those correlations as determined with a  $t$ -test ( $y$  axis). Genes identified in the oligodendrogloma stemness program (this work) are marked in black and are significantly enriched for the GBM stemness genes ( $1.5 \times 10^{-4}$ , hypergeometric test), defined as those with  $P < 0.05$  and an average correlation above 0.1. **d**, Preferential expression of the oligodendrogloma stemness program in neurons but not in OPCs. Genes expressed in the oligodendrogloma single cells were divided into

six bins (bars) based on their relative expression ( $\log_2$ -ratio) in stem-like cells with high PC2/3 and intermediate PC1 scores compared to all other cells. Each panel shows for each bin the average relative expression in each of three normal brain cell types ( $y$  axis) based on data from the Barres laboratory RNA-seq database<sup>9,18</sup>: mice oligodendrocyte progenitor cells (mOPC, top), mouse neurons (mNeurons, middle), and human neurons (hNeurons, bottom). Relative expression of each gene in each cell type was defined as the  $\log_2$ -ratio between the respective cell type divided by the average over AC, OC and neurons. Error bars denote standard error as defined by bootstrapping. Asterisks denote bins with significantly different relative expression (in the respective normal cell type) compared to all genes expressed in oligodendrogloma, based on  $P < 0.001$  (by  $t$ -test) and average expression change of at least 30%. **e**, Correlation with mouse activated NSC program. Shown is the distribution of correlation values ( $x$  axis) of either all genes (grey) or genes from the oligodendrogloma stemness program (black) with the expression program of mice NSC activation states, as previously quantified by 'pseudotime', across single mouse NSCs<sup>19</sup>. The average correlation of the NSC activation program genes with oligodendrogloma stemness genes is significantly higher than with all other genes ( $P = 3 \times 10^{-6}$ ;  $t$ -test). **f**, Correlation with human NPC program. Shown is the distribution of correlation values ( $x$  axis) of either all genes (grey) or genes from the oligodendrogloma stemness program (black) with an expression program of human NPCs identified by PCA (Extended Data Fig. 4). Each gene's correlation to the average expression of the NPC program genes was calculated across single human NPCs. The average correlation with oligodendrogloma stemness genes is significantly higher than with all other genes ( $P = 2 \times 10^{-35}$ ,  $t$ -test).

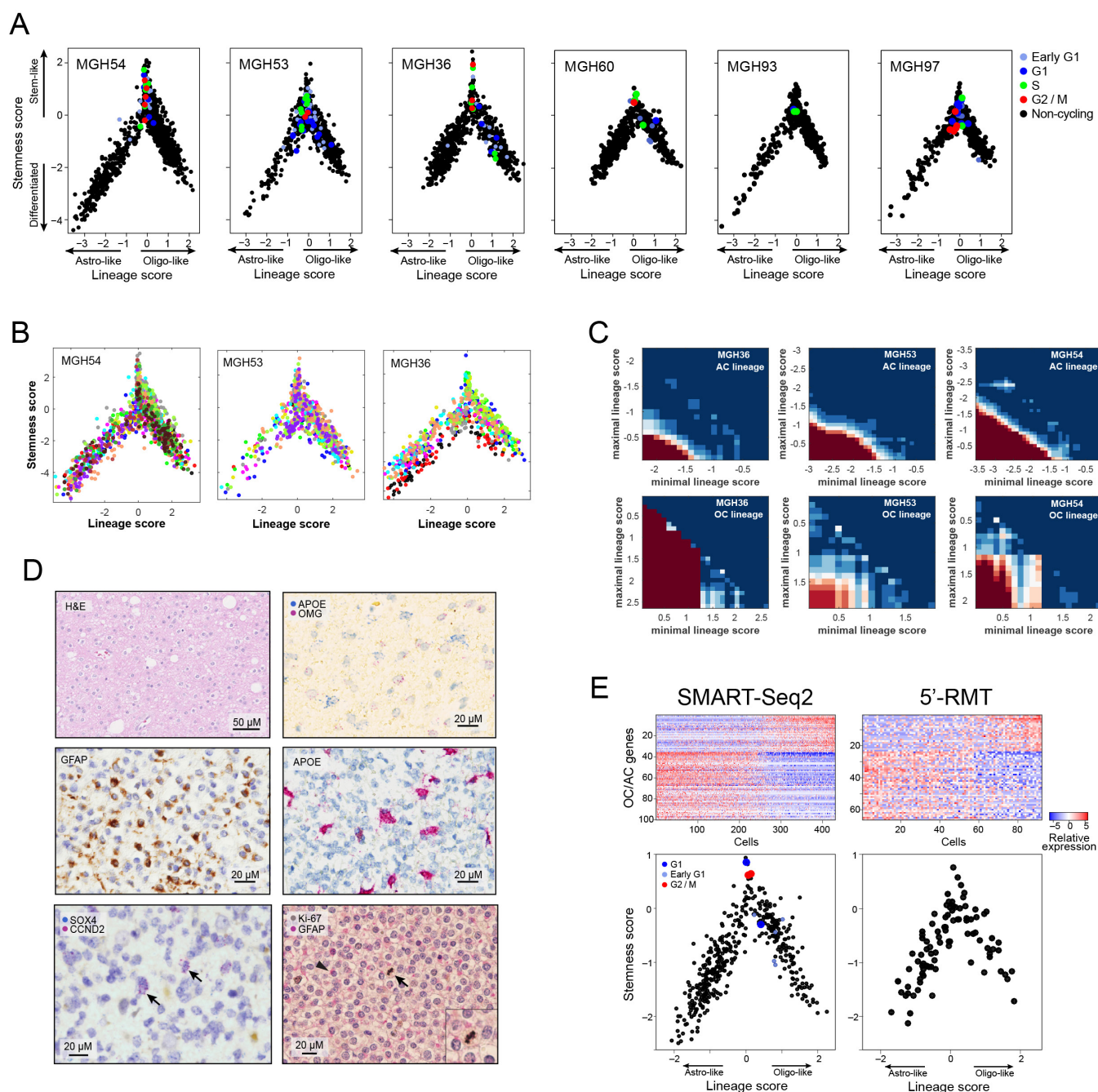


# Extended Data Figure 4 | Analysis of human NPCs.

**a–d**, Differentiation potential of human SVZ NPCs.

Human SVZ NPCs isolated from 19-week-old fetuses form neurospheres in culture (**a**), and can be differentiated to neuronal (neurofilament, **b**), oligodendrocytic (OLIG2, **c**), or astrocytic (GFAP, **d**) lineages *in vitro*. Scale bars, 25  $\mu\text{m}$  (**a**), 10  $\mu\text{m}$  (**b–d**). We note that although OLIG2 can represent different cell types, it is expressed at a low level in the fetal NPCs before differentiation (an average  $\log_2(\text{TPM} + 1)$  of 0.82, compared to a threshold of 4 that we use to define expressed genes in our analysis, and with zero cells with expression above this threshold). Thus, the undifferentiated NPCs do not express OLIG2, and we interpret the expression of OLIG2 as a sign of oligodendroglial lineage differentiation.

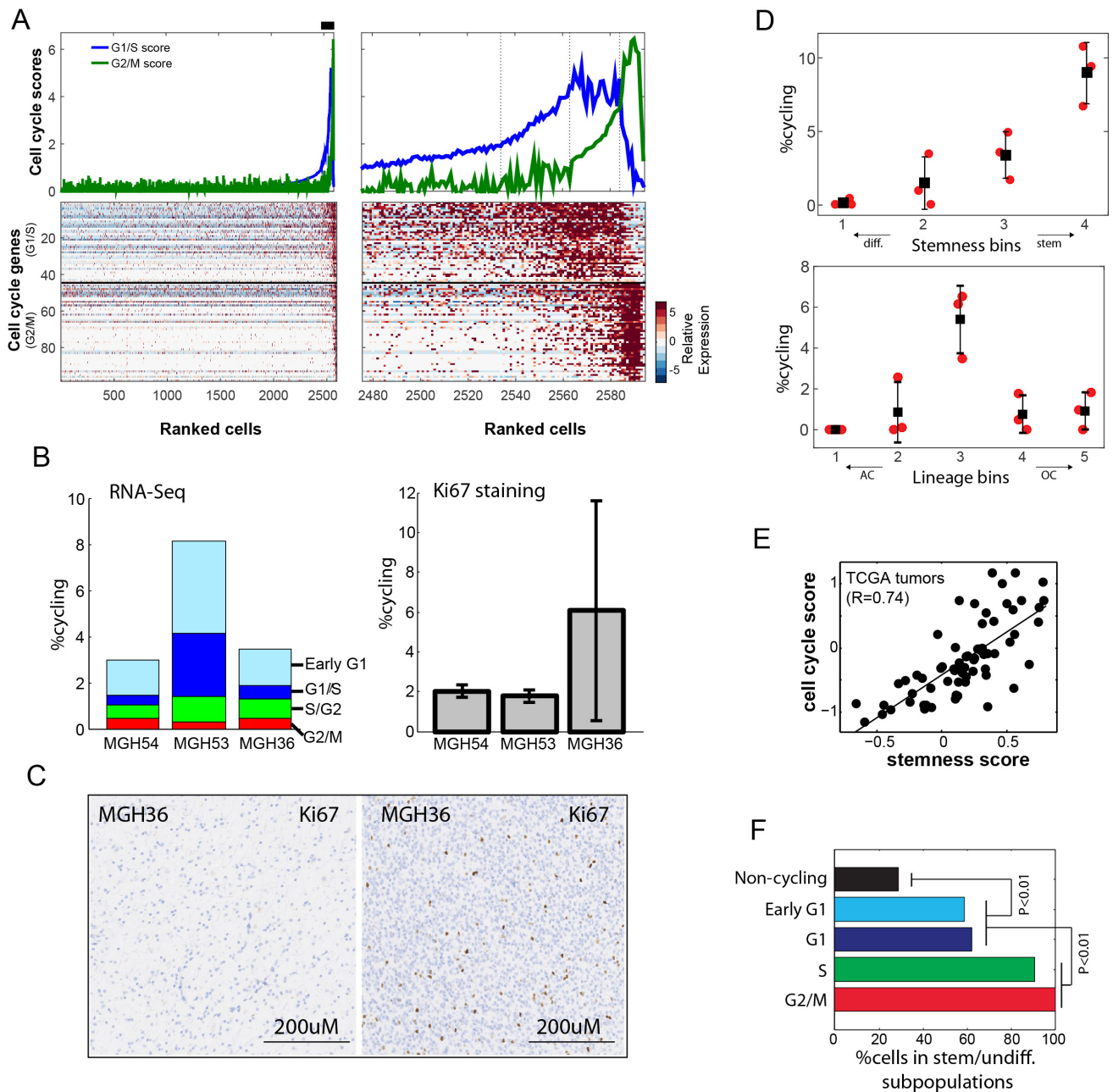
**e, f**, Single-cell RNA-seq analysis of NPCs. **e**, NPCs have an expression program similar to the oligodendrogloma stemness program. Heat map shows the expression of genes (rows) most positively (top) or negatively (bottom) correlated with PC1 of a PCA of RNA-seq profiles for 431 single NPCs, across NPC cells (columns) rank ordered by their PC1 scores. Selected genes are indicated, and a full list of correlated genes for PC1 and PC2 is given in Supplementary Table 2. **f**, NPC cell scores for PC1 (y axis) and PC2 (x axis). PC2 correlated genes are associated with the cell cycle. Cells with the highest PC1 scores tend to be non-cycling (low PC2 score), indicating that while the stemness program is coupled to the cell cycle in oligodendrogloma, it is decoupled from the cell cycle in NPCs.



**Extended Data Figure 5 | Developmental hierarchy in oligodendroglioma.** **a**, Shown are plots as in Fig. 2d for each of the six tumours with cycling cells coloured as in Fig. 3. **b**, Lineage and stemness scores for three tumours with high-depth profiling, coloured based on sequencing batches, demonstrating the lack of considerable batch effects. **c**, For each of the three tumours profiled at high depth (horizontal panels) and for the two lineages (vertical panels), we calculated the significance of co-expression among sets of AC-related (top panels) or OC-related (bottom panels) genes within limited ranges of lineage scores (between the value of the *x* axis and that of the *y* axis). Significance was calculated by comparison of average co-expression to that of 100,000 control gene-sets with similar number of genes and distribution of average expression levels, and is indicated by colour. The significant co-expression patterns within limited ranges of lineage scores suggest that variability of lineage scores in these ranges cannot be driven by noise alone, and implies the existence of multiple states within each lineage, presumably reflecting intermediate differentiation states (see Supplementary Note 2). **d**, Characterization of tumour subpopulations by histopathology and tissue staining. Top/middle panels denote two predominant lineages of AC-like and OC-like cells. Shown are MGH53 with haematoxylin and eosin (H&E, top left), immunohistochemistry for OLIG2 (oligodendrocyte

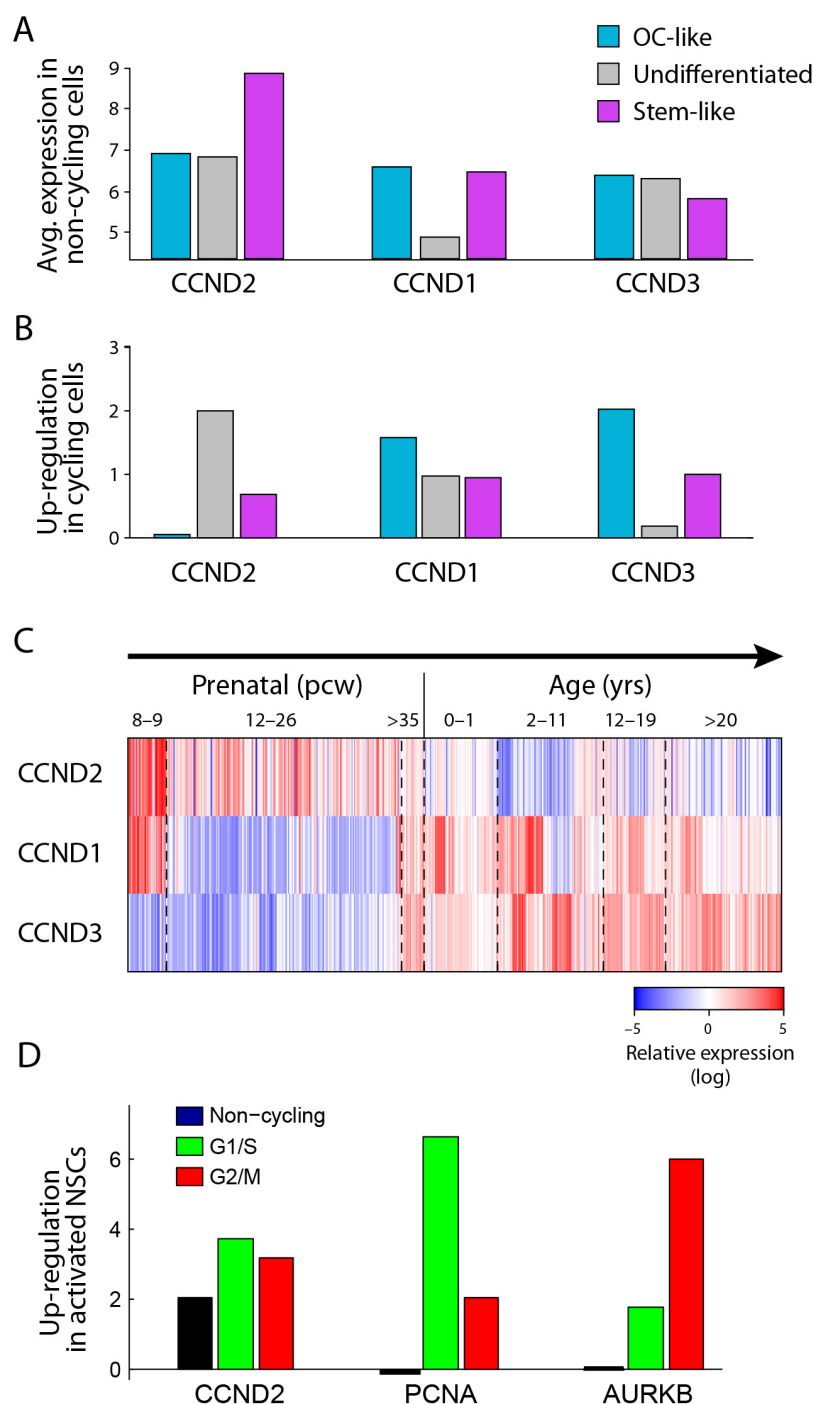
marker, top right) and GFAP (astrocyte marker, middle left), as well as *in situ* RNA hybridization for astrocytic markers ApoE (apolipoprotein E, astrocytic lineage, middle right), with patterns similar to GFAP immunohistochemistry. Bottom panels denote stem-like cells and association with cell cycle. *In situ* RNA hybridization for the stem/progenitor markers SOX4 and CCND2 (bottom left) and the proliferation marker Ki-67 (bottom right) in MGH36 identifies cells positive for both markers (arrows). Immunohistochemistry for GFAP (arrowhead, bottom right) and Ki-67 (arrow, bottom right) shows mutually exclusive expression patterns. **e**, Consistency of MGH60 hierarchy between the full-length SMART-Seq2 protocol used throughout this work (left panels), and an alternative protocol (right panels) in which only the 5'-ends of transcripts are analysed while incorporating random molecular tags (RMTs, also known as unique molecular identifiers, or UMIs) that decrease the biases of PCR amplification. Top panels: PC1 reflects an AC-like and OC-like distinction. Shown are heatmaps of the AC-like and OC-like specific genes (rows, as defined in Supplementary Table 1 and restricted to genes with average expression  $\log_2(\text{TPM} + 1) > 4$  in each data set) with cells ordered by their PC1 score. Bottom, lineage (*x* axis) and stemness (*y* axis) scores (defined as in Fig. 2d).





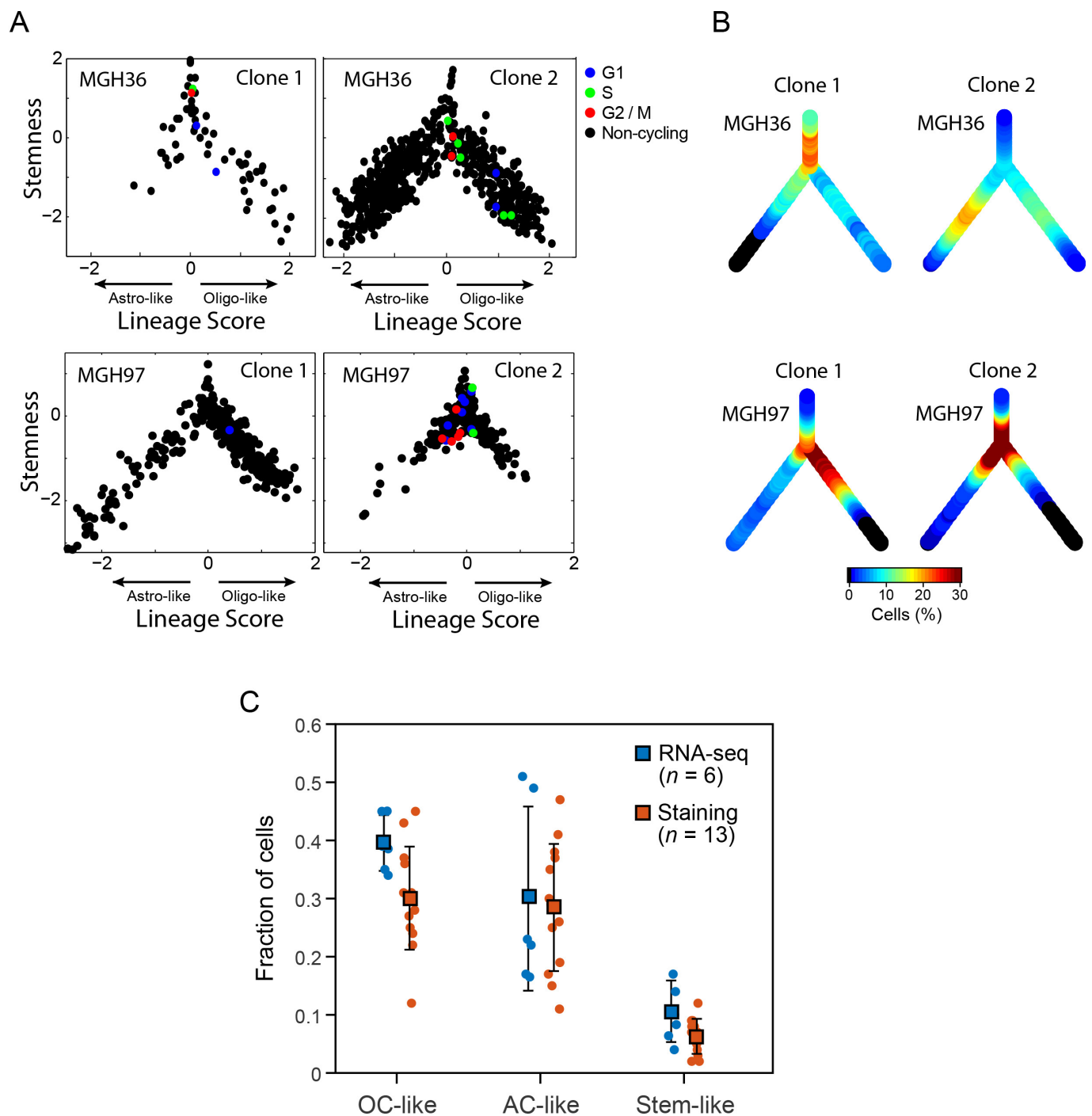
**Extended Data Figure 6 | Cell-cycle analysis.** **a**, High expression of G1/S and G2/M gene sets in a subset of cycling cells. Shown are the average expression (top panels, lines) or the expression of all individual genes (bottom, heat maps) of the G1/S and G2/M gene sets, in all cells ( $n = 2,594$ ) (left) or only among the putative cycling cells ( $n = 119$ ) (right) from the three tumours profiled at high-depth ordered by cell cycle expression. Dashed lines (top right) separates the four inferred phases of cycling cells, corresponding to light blue, blue, green and red in Fig. 3a, respectively. **b**, Estimated fraction of cycling cells (y axis) in each of 3 tumours (x axis) based on single cell RNA-seq (left; different phases marked by colour code as in Fig. 3a) or Ki-67 immunohistochemistry (right). **c**, Variation in cycling cells between regions of the same tumour. Shown is Ki-67 immunohistochemistry in two regions in MGH36. Such regional variability in proliferation complicates direct comparisons as done in **b**. **d**, Cycling cells are enriched in stem-like and undifferentiated cells compared to differentiated cells. Shown is the percentage of cycling cells (y axis) in four bins based on stemness scores (top) or lineage scores (bottom). Black squares and error-bars correspond to the mean and standard deviation of the percentages in the three tumours profiled at high depth (MGH36, MGH53, MGH54), and red circles denote the percentages in individual tumours. Bins in left panel were defined as stemness scores

below  $-1.5$  ( $n = 711$ ), between  $-1.5$  and  $0.5$  ( $n = 1,100$ ), between  $-0.5$  and  $0.5$  ( $n = 939$ ), and above  $0.5$  ( $n = 274$ ), respectively. The first two bins are significantly depleted with cycling cells, while the last two bins are significantly enriched ( $P < 0.05$ , hypergeometric test). Bins in left panel were defined as AC score above 1 ( $n = 503$ ), AC score between 0.5 and 1 ( $n = 1,013$ ), AC and OC scores below 0.5 ( $n = 1,130$ ), OC score between 0.5 and 1 ( $n = 855$ ), and OC score above 1 ( $n = 597$ ), respectively. The third bin is significantly enriched with cycling cells, while the four other bins are significantly depleted ( $P < 0.05$ , hypergeometric test). **e**, Correlation between the average expression of cell cycle (y axis) and that of stemness genes (x axis) across molecularly defined oligodendrogliomas (by IDH mutation, chromosome 1p and 19q co-deletion, and absence of P53 and ATRX mutations) profiled by TCGA ( $n = 69$ ) with bulk RNA-seq. Average expression was defined by centring the log<sub>2</sub>-transformed RSEM gene quantifications. Also shown are the linear least-square regression and Pearson correlation coefficient. **f**, Specific enrichment of S/G2/M cells compared to G1 cells among stem-like or undifferentiated cells. Shown is the proportion (y axis) of each marked category of cells among the stem-like or undifferentiated subpopulations. Significant enrichments are marked ( $P < 0.01$ , hypergeometric test).



**Extended Data Figure 7 | *CCND2* is associated with both cycling and non-cycling stem/progenitor cells.** **a**, *CCND2*, but not *CCND1* or *CCND3*, is upregulated in non-cycling stem-like oligodendrogloma cells. Shown are the average expression levels (y axis, log-scale) of three cyclin D genes (x axis) in non-cycling cells classified as OC-like cells (light blue), undifferentiated cells (grey) and stem-like cells (purple). *CCND2* is approximately fourfold higher in stem-like non-cycling cells than in OC-like and undifferentiated cells ( $P < 0.001$  by permutation test). Conversely, *CCND1* and *CCND3* are expressed at comparable levels in stem-like and OC-like cells. **b**, Upregulation of cyclin D genes in cycling cells compared to non-cycling cells. As in **a** but for up regulation (log<sub>2</sub>-ratio) in cycling cells vs. non-cycling cells. *CCND2* levels further increase in cycling undifferentiated and stem-like cells but not in OC-like cells, whereas *CCND1* and *CCND3* levels increase in OC-like cycling cells more than in undifferentiated and stem-like cycling cells. **c**, Distinct expression patterns of cyclin D genes in human brain development. Shown

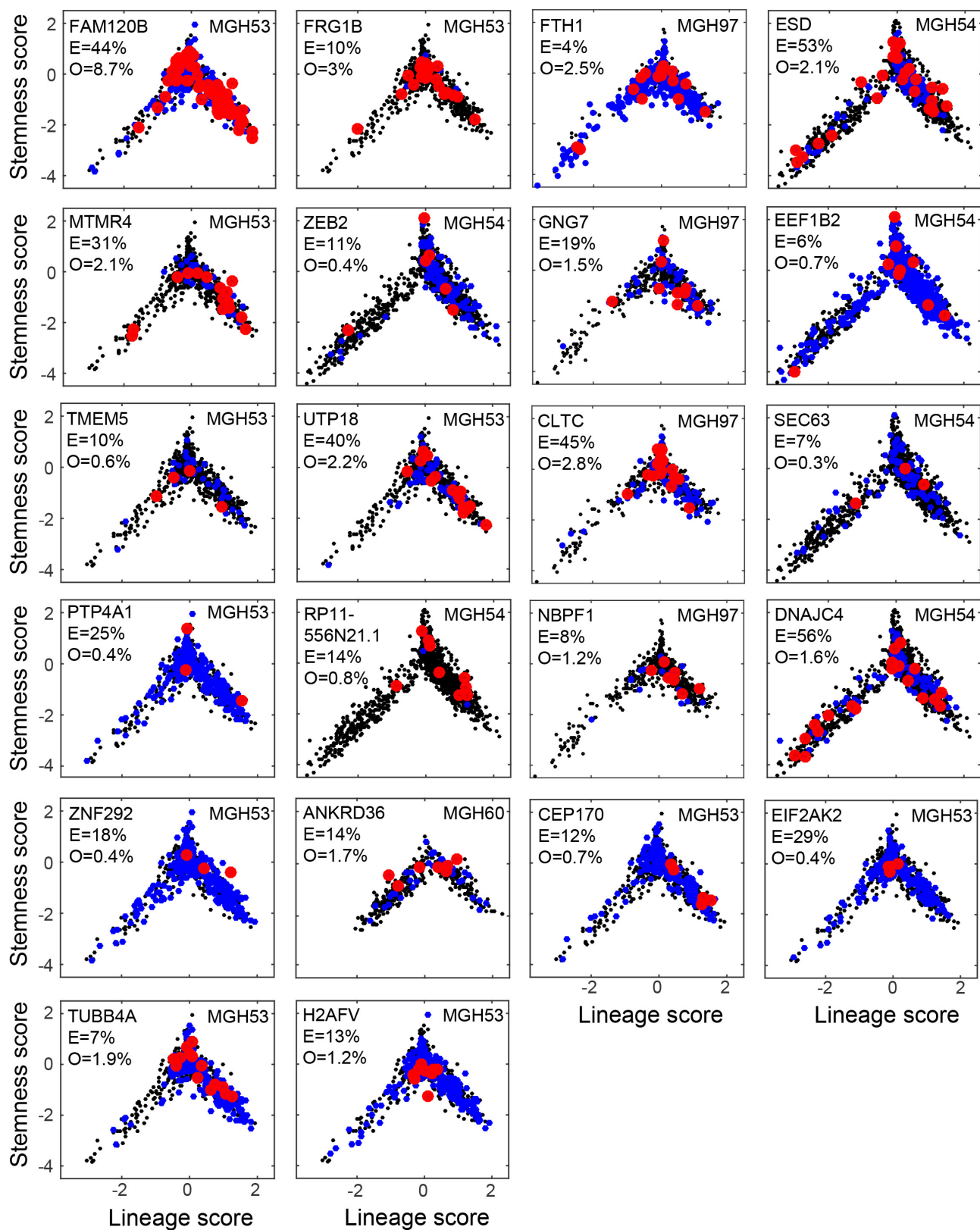
are the expression patterns of three cyclin D genes (rows) in human brain samples at different points in pre- and post-natal development, sorted by age (columns) from the Allen Brain Atlas<sup>15</sup>. *CCND2* is associated with prenatal samples, whereas *CCND1* and *CCND3* are expressed mostly in childhood and adult samples. **d**, *CCND2* is upregulated in activated versus quiescent NSCs<sup>19</sup>, both among cycling and non-cycling cells. Activated NSCs were partitioned into non-cycling cells (black) and cycling cells in the G1/S (green) or G2/M (red) phases (Methods). Expression difference (y axis) for each of three genes (x axis) was quantified for each of these subsets as the log<sub>2</sub>-ratio of the average expression in the respective subset versus the quiescent NSCs, and was significant for each of the three subsets ( $P < 0.05$  by permutation test). Although *CCND2* (left) is induced in both cycling and non-cycling activated NSCs, two canonical cell cycle genes (*PCNA*, middle; and *AURKB*, right) are not induced in non-cycling genes but were induced preferentially in G1/S and G2/M cells, respectively.



**Extended Data Figure 8 | Distribution of cellular states in distinct genetic clones of MGH36 and MGH97.** **a**, Stemness (y axis) and lineage (x axis) score plots for MGH36 (top) and MGH97 (bottom), each separated into clone 1 (left) and clone 2 (right) as determined by CNV analysis (Fig. 1a, b). Cycling cells are coloured as in Fig. 3, with G1/S cells in blue, S/G2 cells in green, and G2/M cells in red. **b**, Colour-coded density of cells across the cellular hierarchy as shown in Fig. 2e, for the two clones

(left: clone 1, right: clone 2) in each of the two tumours (top: MGH36, bottom: MGH97). **c**, The fraction of cells assigned to the different tumour compartments (y axis, Methods) based on either single-cell RNA-seq (blue) or RNA *in situ* hybridization (orange). Circles denote individual tumours; squares denote average of all tumours; error bars denote standard deviation across tumours, showing general agreement between scRNA-Seq and IHC estimates.



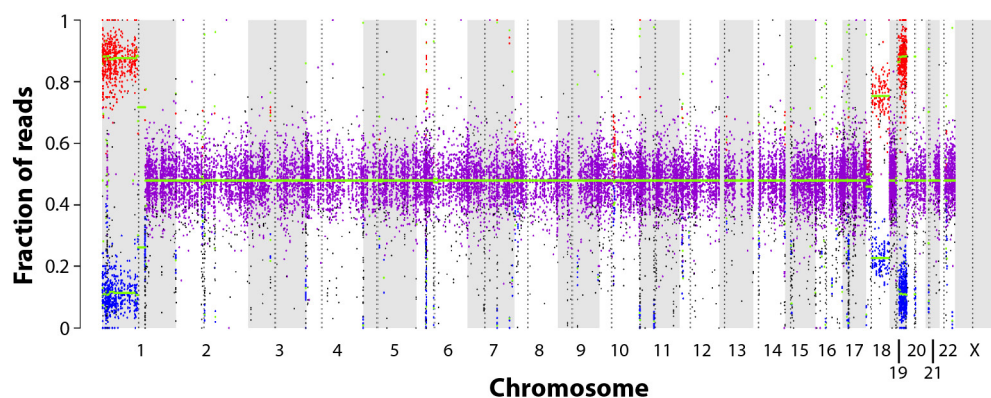


Extended Data Figure 9 | See next page for caption.

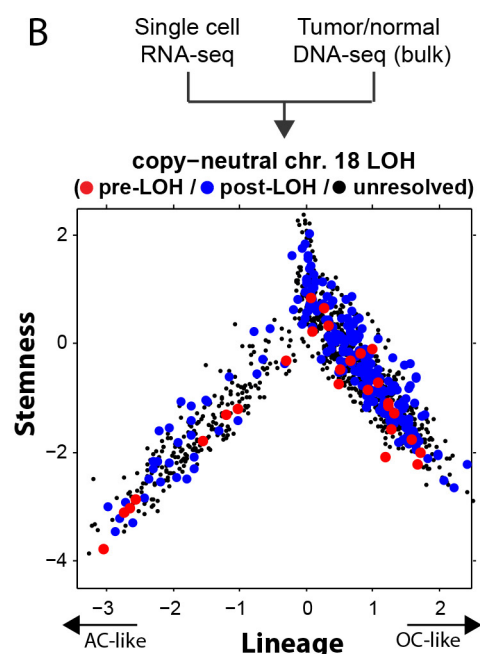
**Extended Data Figure 9 | Subclonal mutations tend to span the cellular hierarchy.** Each panel shows lineage ( $x$  axis) and stemness ( $y$  axis) scores of cells in which we ascertained by single cell RNA-seq a mutant (red), a wild-type (blue) or none (black) of the alleles. Included are mutations for which at least three cells were identified as mutants and that were identified by WES as subclonal (fraction  $< 60\%$ ). The gene names, tumour name, ABSOLUTE-derived fraction of mutant cells (E, expected fraction) and the fraction of cells detected as mutant by RNA-seq (O, observed) are also indicated within each panel. We note that identification of a wild-type allele (blue) does not imply a wild-type cell because mutations may be heterozygous, and thus cells could contain both alleles while only one may be detected by single-cell RNA-seq. The observed fraction of mutations (O) is much lower than expected (E) due to limited coverage of the single-cell RNA-seq data, as well as due to heterozygosity. The vast majority of mutations (20 of 22) are distributed across the hierarchy and span multiple compartments. Two remaining mutations (H2AFV and EIF2AK2) appear more restricted to the ‘undifferentiated’ region

(intermediate lineage and stemness scores), which could reflect our limited detection rate of mutant cells and/or a bias of the mutation to a particular region. To test the significance of potential biases in the distribution of mutations we calculated, for each mutation, a Euclidean distance among all pairs of mutant cells (based on their lineage and stemness scores), and compared the average pairwise distances among mutant cells to that among randomly selected subsets of the same number of cells. None of the mutations were significant with a false discovery rate (FDR) of 0.1, although this could reflect our limited statistical power and we cannot exclude a potential bias. The apparent bias of mutant cells to the OC lineage over the AC lineage (that is, positive versus negative lineage scores) reflects the lower frequencies of AC-like cells compared to OC-like cells in MGH53 and MGH54 (MGH53: 17% AC versus 39% OC; MGH54: 23% AC vs. 45% OC); this bias is also observed for the detection of wild-type alleles (blue) demonstrating that there is no bias against mutation detection in the AC lineage.

A



B



**Extended Data Figure 10 | Loss-of-heterozygosity event in MGH54 reveals two clones that span the cellular hierarchy.** **a**, Chromosome 18 loss of heterozygosity (LOH) in MGH54. Allelic fraction analysis of MGH54 SNPs from WES shows an imbalance (red and blue dots) in the frequency of alternative alleles in chromosome 1p, 19q, as well as chromosome 18, despite the normal copy number at this chromosome (Fig. 1a). This is consistent with an LOH event in which presumably one copy of chromosome 18 was deleted, and the other copy amplified. The weaker imbalance compared to chromosomes 1p and 19q further suggests that this is a subclonal event. **b**, Each of two clones defined by

chromosome 18 LOH status spans the full hierarchy. Shown are the lineage (x axis) and stemness (y axis) scores for each cell from MGH54 ( $n = 1,174$ ) classified as pre-LOH (red), post-LOH (blue) and unresolved (black) based on RNA-seq reads that map to SNPs in the minor (that is, deleted) chromosome. Both the pre- and post-LOH clones span the different tumour subpopulations. Pre-LOH cells were defined as all cells with reads that map to minor alleles in chromosome 18; post-LOH cells were defined as all cells with reads that map to at least five different major alleles, but no reads that map to minor alleles in chromosome 18; all other cells were defined as unresolved.



# CORRECTIONS & AMENDMENTS

---

## CORRIGENDUM

doi:10.1038/nature19781

### **Corrigendum: Human commensals producing a novel antibiotic impair pathogen colonization**

Alexander Zipperer, Martin C. Konnerth, Claudia Laux, Anne Berscheid, Daniela Janek, Christopher Weidenmaier, Marc Burian, Nadine A. Schilling, Christoph Slavetinsky, Matthias Marschal, Matthias Willmann, Hubert Kalbacher, Birgit Schitteck, Heike Brötz-Oosterhelt, Stephanie Grond, Andreas Peschel & Bernhard Krismer

*Nature* **535**, 511–516 (2016); doi:10.1038/nature18634

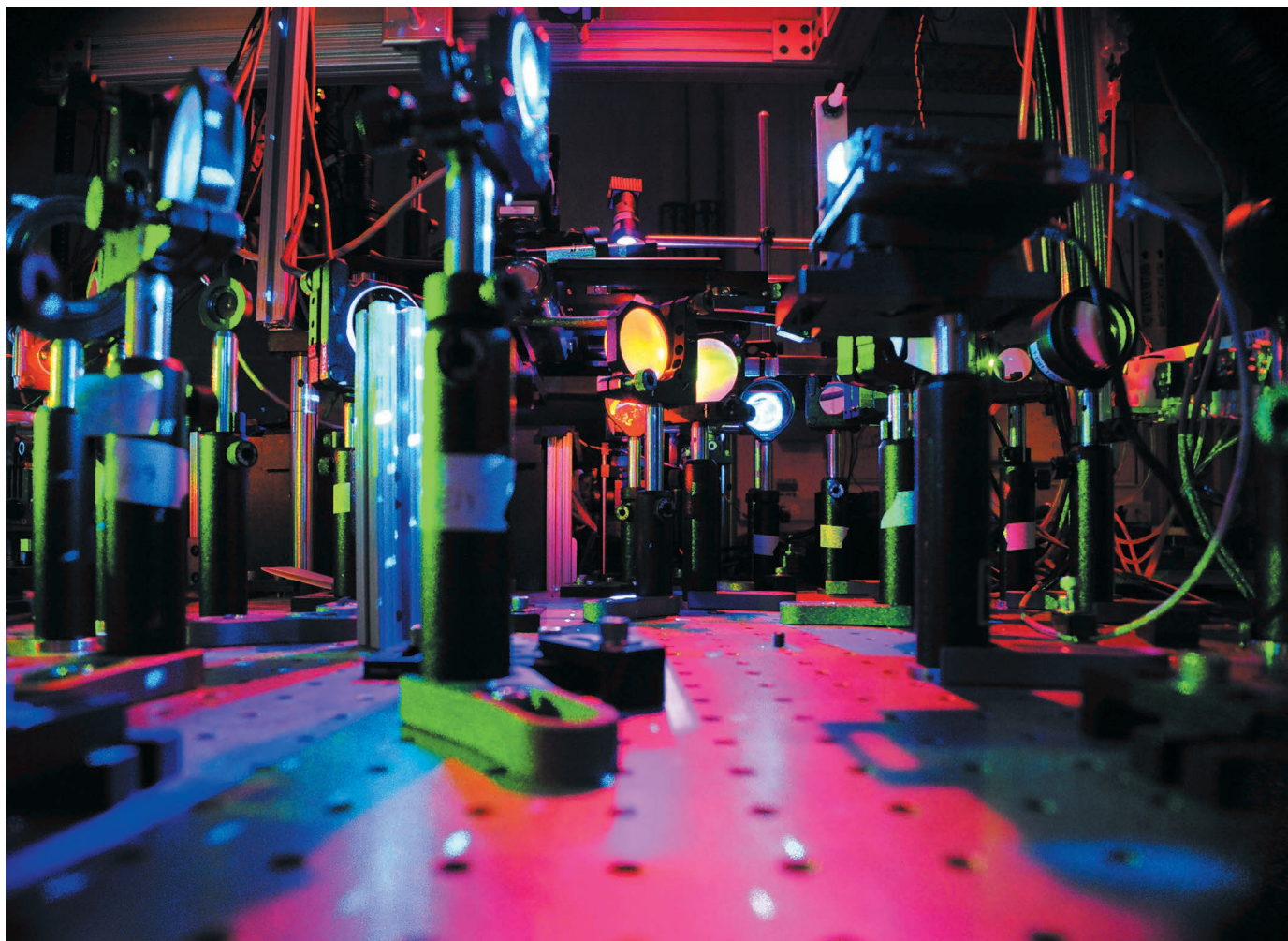
In the interests of transparency, in this Article we wish to amend the competing financial interests statement to read: “Tuebingen University has filed a provisional patent application that covers the compound lugdunin and derivatives thereof, as well as the application of lugdunin-producing bacteria for the prevention of bacterial infections (European patent application number EP 15 160 285.1).” The online versions of the paper have been corrected.

## TECHNOLOGY FEATURE

# THE REAL-TIME TECHNIQUE FOR A LIVING BRAIN

*Neurobiologists are coming up with innovative ways to get high-resolution pictures of the whole brain at work.*

HETIAN



Elaborate arrangements of lasers, mirrors and optics are allowing researchers to probe the brain in ever more creative ways.

BY AMBER DANCE

**R**osa Cossart thinks she knows what a memory looks like. In a study published in *Science* in September, Cossart, a neurobiologist at the Institute of Neurobiology of the Mediterranean in Marseilles, France, opened up mouse brains to visualize their neural activity as the animals raced on treadmills and rested. As the mice

ran, some 50 neurons in their hippocampi fired in sequence, possibly to help the animals measure the distance travelled. Later, when the mice were resting, certain subsets of those neurons turned on again<sup>1</sup>. This reactivation, Cossart suspects, has to do with encoding and retrieving memory — as if the mouse is recalling its earlier exercise.

“The power of imaging is really to be able to see the cells, to see not only the active ones

but also the silent ones and to map them on the anatomical structure of the brain,” she says.

It has not yet provided proof for Cossart’s hypothesis, but the microscope and neural-activity markers behind the techniques represent the very latest in methods to study brain connectivity. In the past, researchers studied just a few neurons at a time using electrodes implanted into the brain. But that gives a fairly crude picture of what is going on, like ►

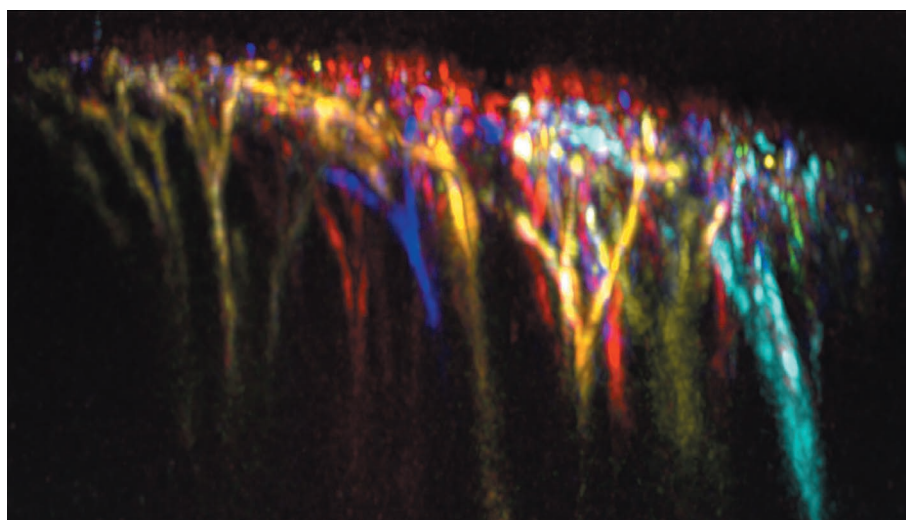
► looking at a monitor with just a couple of functioning pixels, says Rafael Yuste, director of the NeuroTechnology Center at Columbia University in New York City.

But new techniques are fleshing out the picture. Scientists can now watch neurons live and in colour, helping them to work out which cells work together. Methods such as Cossart's zoom in at the microscopic scale to catch individual neurons in the act; others provide a whole-brain, or mesoscopic, view. And although it is possible to perform these experiments with an off-the-shelf microscope, scientists have been customizing them to suit their specific purposes; these devices are in various stages of commercialization.

The field of live-brain imaging is flourishing thanks to innovations such as two-photon microscopy, which allows scientists to image deeper into brain tissue, and indicators that flash as neurons fire; Cossart combined the two in her study.

Major funding initiatives are also pushing the field forward, particularly the US Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative, which aims to improve researchers' ability to map the brain. The US National Institutes of Health has partnered with groups in Canada, Australia and Denmark to co-fund investigators from other countries involved in the BRAIN Initiative. In Japan, the Brain Mapping by Innovative Neurotechnologies for Disease Studies (Brain/MINDS) programme includes funding for projects such as functional magnetic resonance imaging (fMRI) analysis of the marmoset brain.

Nevertheless, the scientists involved in these projects face major challenges. The biggest is the brain matter itself. "Brain tissue has the optical properties of milk," Yuste says. The light waves that microscopists use to visualize neurons tend to bounce off surrounding tissue and scatter in multiple directions. That means that most studies cannot penetrate much more than a millimetre below the brain's surface. But researchers can now use both crude surgical techniques (removing part of the brain to discern what happens underneath, or poking in fibre optics, for example) and tricks of light to



ELIZABETH HILLMAN AND RANDY BRUNO, COLUMBIA UNIV.

New methods such as SCAPE are making it possible to visualize 3D neural activity in animals as they move, as in these apical dendrites in the brain of a living mouse.

sneak their laser beams deeper into the tissue.

Other challenges include the incredible speed at which mammalian neurons communicate as well as how to integrate data all the way from the meso- to the microscale. "The dream is obviously every neuron — every axon, dendrite, synapse — in the whole brain flashing away," says Columbia biomedical engineer Elizabeth Hillman. "We can do it in the fruit-fly brain, and in the zebrafish, just not yet in the mouse."

But despite its limitations, live-mouse-brain imaging is already starting to reveal how neural connections can be silenced, or regrow, in studies of brain disease and ageing.

## OF CALCIUM AND CIRCULATION

Take the work that has been done on stroke at the mesoscale, for instance. Blood clots in the brain damage neurons and thus the routes of neural communication. The damage can easily be seen in people: fMRI has shown that stroke affects the flow of blood between mirror-image parts of the two hemispheres, a cross-talk that is crucial for activities such as coordinated movement. But probing the details of stroke is difficult to do in people, so researchers including neurologist Jin-Moo Lee at Washington

University School of Medicine in St. Louis, Missouri, are keen to use mice as a model to study the disease and possible treatments.

However, mouse brains are so small that fMRI signals get lost in the noise, so Lee had to turn to a different technique to track blood flow. His colleague Joe Culver, a biomedical engineer, introduced him to a technique called optical intrinsic signal imaging (OIS), which picks up colour alterations that are linked to changes in blood oxygen levels. Oxygenated blood is reddish and deoxygenated blood bluish, and the different colours can be detected through the thin skull of a mouse using fairly basic scientific equipment, or even a wearable consumer camera known as a GoPro. Well-oxygenated areas are likely to be more active than others. To study neural connectivity, Culver and his colleagues zoomed out to look at the entire cortex, and presumed that highly oxygenated spots that are flashing in sync are likely to be connected. He calls the new method "functional connectivity optical intrinsic signal imaging", or fcOIS<sup>2</sup>.

In 2014, Culver and Lee used this technique to show that strokes in mice affect connections between mirror-image parts of the two hemispheres, just as they do in people<sup>3</sup>. Culver has also applied fcOIS to a mouse model of Alzheimer's disease and found that cross-hemisphere communication not only drops, but is also correlated with plaque deposition and with ageing. The loss of connectivity seems to happen first, foreshadowing which areas might be vulnerable to plaque accumulation<sup>4</sup>.

The technique provides a good first-pass screen for changes in connectivity, Culver says, because it works on any mouse; some markers, including those used by Cossart, need to be genetically engineered into mouse neurons. Yet it's still only a surrogate marker of brain activity. A step closer is calcium indicators. When neurons receive a signal, an electrical current passes through them. That depolarizes

## AN IMAGING PALETTE

A selection of the genetically encoded neural-activity indicators used in live-brain imaging.

Indicator	Mode	Excitation (nm)	Emission (nm)	Addgene ID
<b>Calcium</b>				
GCaMP5G	2-photon	900–1000 (near infrared)	509 (green)	31788
GCaMP5G	1-photon	488 (blue)	509 (green)	31788
GCaMP6f	2-photon	900–1000 (near infrared)	509 (green)	40755
GCaMP6f	1-photon	488 (blue)	509 (green)	40755
<b>Voltage</b>				
QuasAr2-mRuby2	1-photon	561 (yellow)	600 (red)	59174
QuasAr2-Citrine	1-photon	488 (blue)	529 (yellow)	59172
QuasAr2-mOrange2	1-photon	532 (green)	565 (orange)	59173



the plasma membrane and opens ion channels, allowing calcium to flood into the cell. Indicators in the cytoplasm change shape and fluoresce when calcium flows in, providing more-immediate visual feedback. Among the most popular calcium indicators are the GCaMP proteins (see 'An imaging palette') developed at the Janelia Research Campus in Ashburn, Virginia, which are now in their sixth generation. "We're switching everything we can over to calcium," says Culver.

### FLASHING QUASARS

Calcium indicators have become the workhorses of live-brain microscopy. Scientists can see every neuron — at least in the plane imaged by their microscope — and follow their activity over time. For this kind of focus, scientists often use two-photon microscopy. In standard microscopy, a fluorophore is excited by just one packet of light, so any fluorophore that receives a packet will light up, even if it is outside the focal plane. In two-photon microscopy, scientists use a longer-wavelength laser, so the fluorophore must absorb two photons simultaneously to fluoresce. Because the chance of two photons hitting the same spot is high only at the laser's focal point, the signal is effectively limited to the focal plane. As an added bonus, the longer-wavelength, lower-energy light can penetrate deeper into the tissue. By scanning the laser across the brain, microscopists can build up a high-resolution picture of the brain at a depth of up to one millimetre, Yuste says.

Yet calcium indicators are still only a proxy for the electrical spikes that mediate neural signalling. And they're relatively slow to reflect neural communication — "the smeared-out remnant of a spike", says David Kleinfeld, a neurophysicist at the University of California, San Diego. It takes about 100 milliseconds after the membrane depolarizes for the calcium to bind to the indicator and cause it to change shape and fluoresce, estimates Karel Svoboda, a neurobiologist and biophysicist at Janelia. It also takes half a second or so for the fluorescence signal to decay back to the unlit state, so two or three electrical impulses, or 'action

potentials', could pass in the time that the calcium system can indicate only one. "You probably miss stuff," Svoboda says.

The seventh generation of GCaMP indicators, anticipated within the year, should improve the response speed by an order of magnitude, as well as boost its sensitivity, Svoboda says. But no calcium indicator will ever measure action potentials with the same speed and range as electrodes, he says.

Adam E. Cohen, a biophysicist at Harvard University in Cambridge, Massachusetts, is pursuing a faster type of visual indicator — one that fluoresces as the membrane depolarizes.

The sensors — called genetically encoded voltage indicators (GEVIs) — are based on a protein that allows a Dead Sea microbe to harvest solar energy. Luckily for Cohen and his collaborators, that protein — called Archaeorhodopsin — also fluoresced in response to

*"The dream is obviously every neuron — every axon, dendrite, synapse — in the whole brain flashing away."*

changes in membrane voltage. And Cohen's team was able to enhance the mechanism to create a pair of GEVIs called QuasArs. But the QuasArs were fairly dim, so the group

fused them to brighter fluorophores to create pairs that perform fluorescence resonance energy transfer (FRET), with one fluorophore influencing the emission of the other<sup>5</sup>.

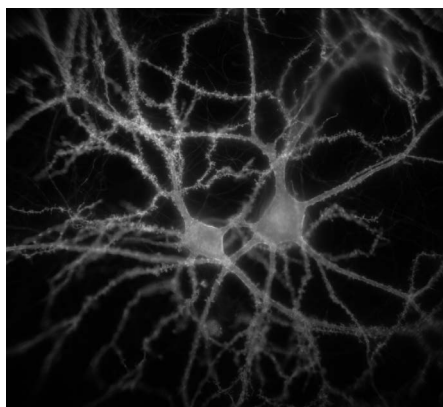
Cohen's team used these indicators, which are available in a handful of colours (see 'An imaging palette'), to monitor spontaneous and induced voltages in cultured rat neurons. Yet, like calcium indicators, GEVIs are neither fast enough nor bright enough to faithfully report neural firing, Cohen acknowledges. And the available colours are restrictive: ideally, he says, GEVIs should fluoresce in the far-red part of the light spectrum, leaving the rest of the rainbow open for the use of other proteins. In particular, neuroscientists often use light-activated proteins to control neuronal activity, using a technique called optogenetics.

### ADVANCING MICROSCOPY

Other researchers are focusing on the microscopy itself, and specifically on imaging in 3D. Because the neurons that work together are not conveniently organized in a single plane, the scanning process must be able to keep pace with signalling across the volume of the brain. Ten frames or 'volumes' per second is a good benchmark, says Fritjof Helmchen, who co-directs the Brain Research Institute at the University of Zurich in Switzerland. "This is one of the clocks the brain is working on" — millisecond resolution would be even better, he adds.

That means that microscope designers must minimize the moving parts that slow things down, says Diego Restrepo, co-director of the Center for Neuroscience at the University

SAMOUIL FARHI



Microscale images can reveal neural activity in a rat hippocampus.



Two-photon microscopy can image cells as deep as 1 mm. These cortical neurons were imaged at 200  $\mu\text{m}$ .

of Colorado Anschutz Medical Campus in Aurora. He and his collaborators have eliminated the up-and-down motion required for focusing by using a liquid objective lens that is controlled by electric field. “When you throw oil on water, you form a lens,” explains Restrepo. By making the lens very small, he and his colleagues have managed to make it very stable, so that it doesn’t bobble about as an animal moves. And they can change the lens’ shape and focal plane by altering the electrical field. Restrepo’s team has used this lens in combination with a confocal microscope and a fibre-optic system to image brain slices<sup>6</sup>, and now plan to attach the device to a mouse’s head.

At University College London, neuroscientist Angus Silver found a way to accelerate the focus changes while imaging across multiple focal planes. He uses an acousto-optic lens that transmits megahertz sound waves through tellurium dioxide crystals to focus the laser beam. “The limitation to speed is the speed of sound across crystal, basically,” Silver says. The technique still isn’t ideal for quickly imaging every neuron in a volume, he says, but it can move from one region to the next in about 25 microseconds<sup>7</sup>. That makes it useful for viewing all of a sparse population, such as inhibitory interneurons in a volume of brain, he suggests.

Another solution to quickly sampling different depths is a modification of light-sheet microscopy, which typically involves moving multiple lenses to continually refocus a sheet

of light. The technique can image one or two volumes per second, Hillman estimates. But by turning the sheet on an angle and using a single mirror to sweep it across the volume of interest, Hillman’s group achieved a rate of 20 times per second. Hillman calls the technique swept confocally aligned planar excitation, or SCAPE, and her team has used it to visualize dozens of distinct firing patterns in the brains of awake mice<sup>8</sup>. The technology has been licensed to Leica Microsystems in Wetzlar, Germany.

Yuste’s group offers yet another option. It uses a spatial light modulator, which splits the laser beam into many beamlets, each of which is aimed at a different part of the tissue. “Imagine a comb of light that’s hitting the sample,” Yuste explains. The microscope picks up any light that comes back, so it can capture multiple planes at once<sup>9</sup>. It can collect about ten sets of images per second, and the researchers are already speeding that up, Yuste says. Yuste has licensed the technology to Bruker in Billerica, Massachusetts, and Olympus in Tokyo, and is contemplating starting his own company.

#### ZOOMING IN, ZOOMING OUT

Most 2D and 3D techniques remain hampered by how the brain scatters light, but scientists have ways of circumventing that limitation, too. At Cornell University in Ithaca, New York, applied physicist Chris Xu and his colleagues reasoned that if two photons could push the imaging depth to a millimetre or so, then three should go even deeper. Indeed, Xu’s three-photon imaging can reach two or three

times further down than two-photon imaging can, he says, although the limits depend on the properties of the tissue being imaged. His group managed to use the technique to image the mouse hippocampus, without removing any of the cortex above<sup>10</sup>.

Xu’s team still can’t penetrate all the way through the brain — “We’re literally still scratching the surface,” he acknowledges — but there’s plenty of room for improvement, he says.

There’s also room to develop live-brain imaging in other ways. A number of researchers, including Kleinfeld and Svoboda, have devised systems that combine the wide mesoscopic field of view with the single-cell resolution achieved by two-photon imaging, allowing them to zoom out on much of the brain or zoom in, Google Earth-style, on individual neurons<sup>11,12</sup>. Kleinfeld’s field-of-view covers an 8×10 millimetre section of cortex; Svoboda’s group can manage a cylinder of brain about 5 millimetres in diameter and 1 millimetre deep, and that’s about 25 times the typical field-of-view in two-photon microscopy, he says. Svoboda has now trained several labs to build their own versions of his microscope, and licensed the technology to Thorlabs in Newton, New Jersey.

Ultimately, these diverse technologies could realize Yuste’s dream for neuroscience: to “crack the code” that links neural firing patterns with behaviour and sensation. The technology can’t yet be used to look at and interpret the activity in a mouse’s visual cortex, for instance, but it has certainly added plenty of pixels to the screen. ■

*Amber Dance is a freelance science writer in Los Angeles, California.*

1. Malvache, A., Reichinnek, S., Villette, V., Haimerl, C. & Cossart, R. *Science* **353**, 1280–1283 (2016).
2. White, B. R. *et al.* *PLoS ONE* **6**, e16322 (2011).
3. Bauer, A. Q. *et al.* *NeuroImage* **99**, 388–401 (2014).
4. Bero, A. W. *et al.* *J. Neurosci.* **32**, 4334–4340 (2012).
5. Zou, P. *et al.* *Nature Commun.* **5**, 4625 (2014).
6. Ozbay, B. N. *et al.* *Opt. Lett.* **40**, 2553–2556 (2015).
7. Naga Srinivas Nadella, K. M. *et al.* *Nature Methods* <http://dx.doi.org/10.1038/nmeth.4033> (2016).
8. Bouchard, M. B. *et al.* *Nature Photon.* **9**, 113–119 (2015).
9. Yang, W. *et al.* *Neuron* **89**, 269–284 (2016).
10. Horton, N. G. *et al.* *Nature Photon.* **7**, 205–209 (2013).
11. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. *eLife* **5**, e14472 (2016).
12. Tsai, P. S. *et al.* *Opt. Express* **23**, 13833–13847 (2015).

#### CORRECTION

The Technology Feature ‘The dark side of the human genome’ (*Nature* **538**, 275–277; 2016) implied that Ran Elkon was solely responsible for performing the first screen using the advanced editing system. In fact, he was part of the team headed by Reuven Agami.

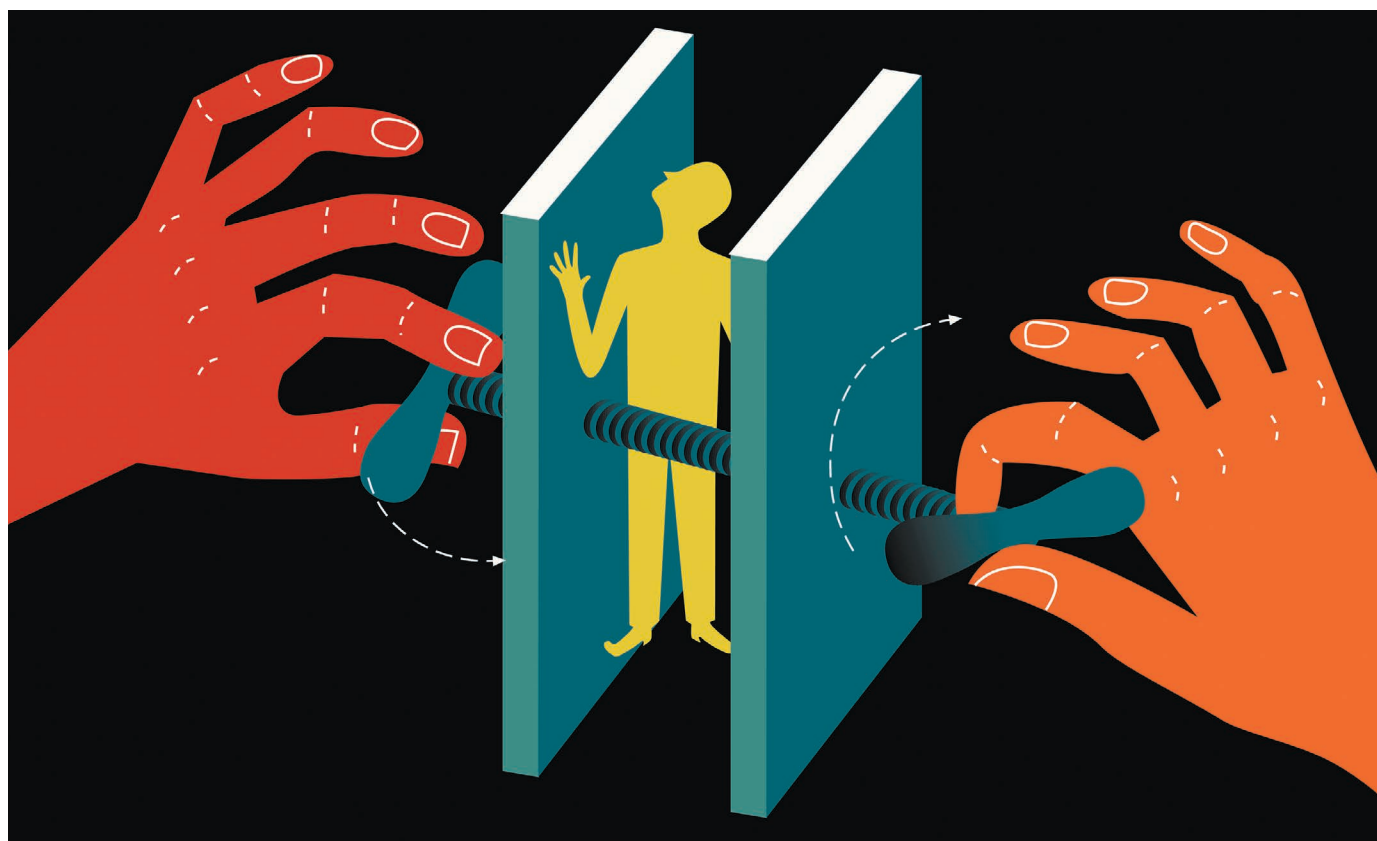
# CAREERS

**FUNDING** European universities call for more money **p.321**

**MINORITIES** US agency aims to boost diverse faculty numbers **p.321**

**NATUREJOBS** For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)

ANDY BAKER/GETTY



MENTAL HEALTH

## Caught in a trap

*The pressures of a scientific career can take their toll on people's ability to cope.*

BY EMILY SOHN

Rachel was working towards her PhD in 2008 when depression first threatened to derail her career. The psychologist was clocking 14- to 16-hour days to launch a study and was overseeing 12 research assistants. In her spare time, she was planning her wedding.

For three weeks, she barely slept and subsisted on sweets and energy bars. Then came a crying spell that started at a sad theatre performance and turned into three days of uncontrollable sobbing. "I was sitting on my couch, staring at the wall and crying," says Rachel, who requested that her name be changed for this article.

The seeds of Rachel's depression pre-dated her PhD. "You are bombarded with messages before you even apply for PhD programmes — that it's hard to get in, that 50% don't finish, that it's hard to get postdocs, that it's impossible to get grants," she says. "At the same time, you are surrounded by people who have PhDs. If you already have a tendency toward perfectionism or self-doubt, it feeds that really nicely."

She returned to work after a two-week break that included therapy and a prescription for antidepressants. After completing her dissertation, she landed an assistant-professor post at a university in New York.

"If you had asked me at the time, I would have said, 'Oh no, I'm never finishing. I'm

leaving academia,'" she says about her break. When she shared her feelings with her father, who also holds a PhD, he told her, "Welcome to science."

Depression and anxiety are widespread, including among scientists, who often face intense pressure to work long hours, publish in high-profile journals, win grants to support themselves and others and rebound after repeated rejections. Depression affects 350 million people around the world and is the leading cause of disability globally, according to the World Health Organization (WHO).

Each year in the United States, almost 7% of adults, or an estimated 16 million people, have at least one major depressive episode. To be diagnosed with the condition, people ►



► must experience a minimum of five specific symptoms that impair functioning or cause significant distress almost every day for at least two weeks. One of those must be a persistently low mood (marked perhaps by a sense of emptiness or hopelessness) or a loss of pleasure or interest in almost all activities. But they might also include recurrent thoughts of death or significant weight loss, for example.

The chance that a person will develop depression during their lifetime varies by nation — in the United States, it is 17%. Depression frequently goes hand-in-hand with anxieties or worries that are so excessive that they interfere with life. Such anxiety affects about 18% of US adults, or 40 million people. More than 25% of Europeans have some form of anxiety or depression, according to the WHO.

Research on the prevalence of anxiety and depression specifically in scientists is scarce. But risks are probably comparable to those for the rest of the population, says Jerald Kay, a psychiatrist at Wright State University in Dayton, Ohio, and co-editor of *Mental Health Care in the College Community* (Wiley, 2010), whose practice has focused on physicians, university students and faculty members, including researchers.

Some situations are more likely to lead to anxiety and depression. Graduate studies can be particularly tough, because students suddenly face high expectations and low salaries and find that their fates lie in the hands of advisers, who can even live in another country.

In a 2014 survey of 790 graduate students at the University of California, Berkeley, almost half of PhD students met the criteria for depression, including up to 46% of those in a category that included biological and physical sciences. The stigma associated with mental

illness can make many researchers cautious about revealing their struggles to anyone, let alone to superiors. Still, experts say that it's important to seek help, at least from professionals. "If you feel like you're the only one struggling with a problem, you think, 'It's my personal problem, it's my fault,'" says Joeri Tjink, a psychiatrist at the VU University Medical Center in Amsterdam, "not a fault of the system."

Most universities have mental-health services that offer confidential help, although many researchers who have weathered the storm of depression list superiors and colleagues as sources of support. In a high-pressure career that values prestige but is rife with criticism, it helps to know that you're not alone (see 'Dealing with depression').

### HELPING HANDS

Shweta Ramdas struggled quietly with depression for several years. Originally from India and a graduate of the National University of Singapore, Ramdas found winters increasingly intolerable when she moved to the University of Michigan, Ann Arbor, to do a PhD in bioinformatics.

On some days she would just stare at her computer, taking weeks to finish one-day tasks. Uninterested in food and tired of feeling unhappy, she thought about quitting.

Ramdas finally told her department chair, who said that others had faced similar problems. Soon colleagues were telling her their own stories. "These were amazing people who I really admire," she says. "And they did not seem from the outside to be depressed."

Ramdas spent eight months with family in India, where she saw a therapist. Her professors in Michigan told her to take as much time as she needed. She returned this autumn

determined to set limits on work.

Now she is talking to her supervisors about helping others who struggle with depression. "I feel like graduate schools can do a better job about getting this into the open a lot more," Ramdas says. "I could have handled it better if I had known it wasn't just me."

Elizabeth Droge-Young experienced her first depressive episode in early 2012 while studying the mating systems of promiscuous beetles for a PhD at Syracuse University in New York. At first, she would sit on the couch in her pyjamas playing video games, watching films or listening to sad music and questioning the meaning of life.

As time passed, she became unable to get to the lab or take showers. In 2014, she spent ten days in a hospital almost within view of her lab. "As my life started falling apart, science was really the last thing I held on to," says Droge-Young, who earned her degree this year and is now a science writer. "It kept me going for quite some time, until my depression became too big an issue."

The transition to graduate studies can be jarring, says Matthew Wilkins, an evolutionary biologist at the University of Nebraska-Lincoln. When he started his PhD programme at the University of Colorado Boulder in 2008, he was caught unawares by the sudden expectation to be self-driven and by the ruthlessness of a career in science, in which successes are often punctuated by rejections — for grants, publications and jobs.

That endless need for external validation can foster anxiety, adds Wilkins, who once received a disappointing score on a high-stakes exam — a week before landing a prestigious fellowship. "In academia, success is not guaranteed," he says. "You recognize that it's going to be hard. I don't think you recognize that it's going to take a psychological toll on you."

Tjink chose to study the psychological effects of publication pressure, partly because of conversations with therapists who treat scientists. "They feel so pressured," the therapists told him. "They are exhausted. They are suspicious of people stealing their ideas. Or they feel that colleagues want their positions."

He wanted to make the problem more visible. In a survey of more than 400 medical academics in the Netherlands, published in 2013, Tjink reported that nearly 25% met the criteria for burnout, which is defined as emotional exhaustion (J. K. Tjink et al. *PLoS ONE* 8, e73381; 2013).

Some scientists struggle when they venture beyond conventional research tracks. Paul Andrews, an evolutionary psychologist at McMaster University in Hamilton, Canada, says that when he was a postdoc, he thought his prospects were good because he had published well-cited papers in good journals. But he couldn't find a position that allowed him to specialize in the biological basis of depression. Frustrated, he lost his motivation to eat,

## DEALING WITH DEPRESSION

### Resources to help those who are struggling

Many scientists are likely to encounter a colleague who is struggling. Experts recommend assuring the person that they are not alone and that it's OK to ask for help. Checking in by calling or texting, and listening to what they have to say can help the person to feel cared about and less isolated.

Don't tell someone to snap out of it or cheer up. Don't offer simple-sounding solutions, such as yoga or a new diet. It can be insulting to imply that distractions or easy fixes will solve depression. Those kinds of words can also make someone feel like a failure for being unable to get better alone. And don't ask if they are all better now. Healing is a long process and does not necessarily have an end point.

For anyone who is struggling themselves, or would like to pass on information about resources, some include:

- Anxiety and Depression Association of America: [go.nature.com/2fibv1r](http://go.nature.com/2fibv1r)
- US National Institute of Mental Health: [go.nature.com/2f0vxcw](http://go.nature.com/2f0vxcw)
- Mind, for better mental health: [go.nature.com/2eg6ffs](http://go.nature.com/2eg6ffs)
- Mental Health Europe: [go.nature.com/2fijxxv](http://go.nature.com/2fijxxv)
- World Health Organization on depression: [go.nature.com/2f0ntsqq](http://go.nature.com/2f0ntsqq)
- World Health Organization on mental health: [go.nature.com/2f7fmb2](http://go.nature.com/2f7fmb2)
- The UK Student Mental Health Charity's Guide for Friends: [go.nature.com/2euxzdd](http://go.nature.com/2euxzdd) **E.S.**



ANDY BAKER/GETTY

sleep and exercise. And he obsessed about whether to try getting one big study into an influential journal or to churn out lots of smaller papers.

Andrews took the riskier first approach, eventually publishing a paper in *Psychological Review* in 2009 (P. W. Andrews and J. A. Thomson Jr *Psychol. Rev.* **116**, 620–654; 2009) that received lots of publicity, including a feature about his work in *The New York Times Magazine* early in 2010. Even after that, he had trouble getting interviews, despite submitting many applications. “I was like, ‘What do I have to do to get a job?’” says Andrews, who was by then desperate, depressed and anxious.

He still has recurring depressive symptoms when he struggles with obstacles at work. His research, which challenges mainstream ideas about depression as a disorder and the role of serotonin, also raises doubts about the value of antidepressants. He has found it hard to get his work published.

### TRICKY TOPICS

Sometimes, research topics can induce dark thoughts, says Alejandro Frid, who started studying endangered deer in Chile in 1990 and later investigated the effects of marine fisheries on predator–prey interactions in Alaska and British Columbia.

The more his research pointed to the damaging implications of climate change, the more angry and nihilistic he became. “By nihilism, I mean that there is no point in caring about the future because there really isn’t one,” says Frid, who is now a science coordinator at the Central Coast Indigenous Resource Alliance near Vancouver, Canada, and wrote the book *A World For My Daughter: An Ecologist’s Search for Optimism* (Caitlin, 2015). “It’s all doom. Humans are destructive and we don’t know any better. There’s no real vision worth living for.”

Depression has become a recurring topic

of conversation among environmentalists. Frid has cultivated hopefulness by focusing on ecological resilience and the human capacity for problem solving. He is also working to influence policies and human behaviours that affect the environment — approaches that helped to reshape his career and brighten his outlook.

Scientists can be wary of admitting to symptoms of depression or anxiety. But researchers who confide in colleagues say they’re often surprised by the support they receive.

Like Ramdas, Droge-Young found that when she overcame her fear of telling lab colleagues about her mental illness, others told similar stories. One had spent time in the same hospital. Faculty members contributed to a Kickstarter funding campaign for two art shows that describe her experience of depression and self-harm. Her adviser and his family attended. “People are really caring,” she says, “if you open up with your vulnerabilities.”

Not everyone feels comfortable talking to their superiors, but they should really see a professional as early as possible, says Kay. “If you think you’re struggling, that’s a good reason to get help,” he says. And if one therapist doesn’t seem like a good fit, find another, Droge-Young advises.

Allowing time for outside interests can help to alleviate work-related anxiety. For Droge-Young, that means getting outside or throwing Oscars-watching parties. Wilkins likes to rock-climb, play football and run. He also recommends developing short-term projects with quick deadlines that aren’t work-related. This year, he and a friend started entering — and winning — film competitions.

Recognizing that it is normal and even helpful to feel down when faced with complex problems may also help scientists to cope, says Andrews. He sees his own bouts of depression not as a sign of a malfunctioning brain, but as a response to important problems. That response helps him to focus. When prompted by social problems or work stresses, he says, the body reallocates energy to the brain. The hypothalamus kicks in, suppressing libido and other physical drives and inducing a fixation on negative thoughts.

Although such rumination is often seen as a bad thing, dwelling on a problem can actually help to solve it because it helps the mind to break it down into smaller components. Just like a scientist does, Andrews says. ■

**Emily Sohn** is a freelance journalist in Minneapolis, Minnesota.

### EUROPE

## Call to keep funding

A group of leading European research universities is calling on the European Commission to restore funds cut from the €80-billion (US\$88-billion), 7-year pan-European research funding scheme Horizon 2020. The 21-member League of European Research Universities (LERU) says in *LERU’s Interim Evaluation of Horizon 2020* that the diversion of billions to support risky research has resulted in less funding and lower application success rates. The low success rates will dissuade talented researchers from applying for Horizon research grants in future, the report authors warn. LERU calls for the commission to increase grant-scheme budgets, and warns that the commission focuses too much on technical innovation and ‘cutting-edge’ research. The authors also suggest that the commission should repeat popular funding calls to offer more chances to apply for grants and provide extra funding to support maternity and leave options at universities. They also call for the creation of large, broad panels to evaluate collaborative proposals more effectively.

### POLICY

## Minorities programme

A US\$5.9-million expansion of a National Science Foundation (NSF) programme aims to develop reliable career pathways in science, technology, engineering and maths (STEM) for people in under-represented minority groups, who comprise 8% of faculty appointments at 4-year US academic institutions. The Alliances for Graduate Education and the Professoriate programme will provide funding to 14 universities to address issues, policies and practices that limit the advancement of people from these groups in academic science. Among the concerns that the NSF wants to address are the difficulties that people from minority groups have in navigating career transitions, such as from graduate student to postdoc and postdoc to faculty. Funded universities will collaborate on improving the advancement and retention of women of colour in faculty posts, increasing the success rate of under-represented-minority postdocs seeking STEM faculty posts and improving the experience of PhD students presenting their dissertations. The NSF offers both a Graduate Research Fellowship and a Postdoctoral Research Fellowship.

# MELISSA

*Lost time.*

BY TROY STIEGLITZ

The receptionist looked up from her desk with a warm smile. “Can I have your name please?”

“Yes ma’am, my name is Captain Carter Harrison. I’m here to visit Mrs Melissa Taylor.”

He was a young man with short-cropped hair and an athletic build. He wore a decorated military uniform. He shifted his weight nervously from one foot to the other and crumpled and folded his hat between his hands.

“Oh, Mr Harrison — Melissa has been expecting you. She’s been asking the staff about you all week. Now don’t take it personally, but she’ll probably forget your visit and will be asking about you all over again tomorrow. Here’s your visitor’s identification,” she said as she handed over a white badge. “Please use the computer to check in and head down the hallway to your left. She’s in room 3417.”

He nodded, took a deep breath and started down the long antiseptic corridor. His palms were clammy and his heart was pounding as he peeked in through the open doorway of the room.

An elderly woman was lying peacefully in bed with her eyes closed. A blue and white knitted blanket was pulled up to her waist. He paused for a moment and then softly knocked three times on the metal door frame. Melissa slowly opened her eyes with a smile and whispered: “Daddy.”

“Hi, sweetie,” he said as he choked back tears and entered the room. She had aged so much since he had seen her a few months before — of course, a few months’ ship time meant that years had passed on Earth.

“How was your trip, Daddy? Did you bring me back anything?” Her eyes were sparkling, but the sentences were strained. She had never resented his absence from

her life — even into old age she had always welcomed him home with a girlish enthusiasm.

“I sure did, kiddo.” He reached into the front pocket of his jacket. “I brought you this coin from the new colony on Stratus 8 and this diamond from the mines on Archibald. The crew was able to get some time away from the ships during our refuelling stops.”

She smiled and nodded towards her dresser. Hundreds of small trinkets, stones

thesis defence and I missed your wedding ceremony...” he trailed off as his bottom lip began to tremble. He bit it and wiped back his tears with the sleeve of his itchy wool uniform. Then he gently took her wrinkled hand with his own.

“I was only supposed to work as a captain for a couple more years, but after your mother died we needed the money, sweetie — and we had just lost so many men in the First Interstellar War...” He stopped and fought back his emotion. “I couldn’t just abandon the Terran Fleet after that. I had — I *have* a duty to our planet. I know that I’ve already told you this a thousand times, sweetie.”

The captain sat and held her hand for a few more minutes in silence. He pulled the knitted blanket up to Melissa’s chest and drew the window blinds.

He knew that this was probably the last time that he would see his daughter alive. It would be nearly five years’ Earth time before

he returned from his next assignment.

“I love you, Melissa.” He gently kissed his daughter on the forehead, put on his hat and quietly left the room.

Outside he hailed one of the waiting air taxis. He paused before pressing the hatch-release button and looked to the sky. Hundreds of air cars swarmed upwards into the clouds to dock with one of the massive ferries floating above the city. His ferry would be leaving the atmosphere soon to rendezvous with his convoy of orbiting starships. He still had another 30 years of planetary service before he retired, he thought. He would lose hundreds of years of Earth time before his service was complete. Each time he returned home he would learn that loved ones had passed away and that new ones had been born. He drew a deep breath, pressed the hatch-release button and climbed inside. ■

*Troy Stieglitz is a Golden Age science-fiction enthusiast and fledgling writer working in the field of environmental chemistry.*



ILLUSTRATION BY JACEY